AD-A254 634

ALGORITHMS FOR SENSOR FUSION
Applications of Distance Measures and Probability of Error Bounds to Distributed Detection Systems

Syracuse University

Pramod K. Varshney and Wael Abdel-Mohsen Hashlamoun

DTIC
S ELECTE
AUG 3 1 1992
A D

92-24004

92 8 28 106

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1991	Final Oct 90 - May 91

4. TITLE AND SUBTITLE
ALGORITHMS FOR SENSOR FUSION, Applications of Distance Measures and Probability of Error Bounds to Distributed Detection Systems

5. FUNDING NUMBERS
C - F30602-89-C-0082
PE - 61102F
PR - 2305
TA - J8
WU - 04

6. AUTHOR(S)
Pramod K. Varshney
Wael Abdel-Mohsen Hashlamoun*

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Syracuse University
Dept of Electrical & Computer Engineering
Syracuse NY 13244-1240

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (OCTS)
Griffiss AFB NY 13441-5700

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
RL-TR-91-423, Vol II
(of two)

11. SUPPLEMENTARY NOTES   (See reverse)
Rome Laboratory Project Engineer: Vincent C. Vannicola/OCTS/(315) 330-4437
*Submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in Electrical Engineering in the Graduate School of Syracuse University

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
In this report, we consider some design and analysis aspects of distributed detection networks. The main focus is on the Bayesian approach to the design of these systems. They present a computationally efficient approach to the design of decentralized Bayesian detection systems. This procedure is based upon an alternate representation of the minimum average cost in terms of a modified form of the Kolmogorov variational distance. They demonstrate the utility of our approach by applying it to the design and performance evaluation of four decentralized detection structures.

14. SUBJECT TERMS
Sensor Fusion, Signal Processing, Algorithms

15. NUMBER OF PAGES
178

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Block 11 (Cont'd)

Prime Contractor: Kaman Sciences Corporation, 1500 Garden of the Gods Road,
                    Colorado Springs CO 80933-7463

Acknowledgement

I wish to express my deep gratitude to Dr. Pramod Varshney for his guidance and encouragement throughout the course of this work. He has been helpful, supportive, and understanding. My experience in being a student of his will not be forgotten. I also wish to thank Dr. H. Wang for being the reader of the dissertation, Dr. D. Weiner, Dr. H. Schwarzlander, Dr. K. Kano, and Dr. K. Mehrotra for their acceptance in being on my examination committee.

My beloved wife Raja needs special thanks. She has been patient and tolerant. She did an outstanding task of keeping quiet the noisy children Anas, Mohammed, and Kotaybah. She is really special.

Finally, I express my thanks to the administration of Birzeit University for allowing me to leave my job and giving me this opportunity.

Table of Contents

CHAPTER 1

INTRODUCTION

1.1 Motivation

In large organizational structures where the decision making process is too compli-
cated to be handled by a single agent (or decision maker) or where agents are distributed
over a wide geographic area, decentralized decision making paradigm is employed fre-
quently. Examples of such decision making organizations occur in financial institutions,
industrial environments, and military surveillance systems etc. In order to make the
group's decision more efficient and to avoid the dominance of an individual member over
the rest, the group is divided into smaller decision making units consisting of one or more
members (experts). Decisions of these smaller units are conveyed to the head of the orga-
nization who is responsible for decision combining and reaching the final decision.

One interesting application of group decision making that has received an increasing
interest in recent years is the design and analysis of distributed sensor networks for signal
detection (decentralized detection systems). This is due to the advantages that these sys-
tems have over their centralized counterparts like reliability, survivability, and shorter de-
cision times. In a decentralized detection system there is a group of remotely located local
detectors that monitor a common phenomenon. These detectors are linked to a primary de-
tector (also known as the global decision maker or data fusion center) through bandlimited
channels. Due to this limitation on the bandwidth, the local detectors are not able to convey
to the primary detector all the information available to them about the phenomenon. In-
stead, they provide the primary detector with a compressed version of their data. The role
of the primary detector is to combine the preliminary decisions of the local detectors
along with any locally received observation to make the final decision. For a given opti-
mality criterion, the design of a decentralized detection system involves specifying both
the local decision rules and the global decision rule. Due to the constraints on the trans-

mission capacity of the communication channels, the system experiences performance degradation and computational difficulty in the design of the optimal system. In the absence of these constraints, the configuration reduces to a centralized system for which a well developed theory exists [1-3]. Because of the full utilization of the raw observations in the centralized system, its performance is used as a benchmark for comparing the performance of different decentralized detection network structures based on different design criteria. In general, the design of an optimum decentralized detection configuration is a computationally difficult task [4]. Therefore, rather than resorting to the optimum design criterion, suboptimum design criteria may be used instead. A design criterion A is said to be "better" than a design criterion B if the performance of the system when criterion A is employed is closer to the performance of the centralized system than the performance of the system when criterion B is employed.

Figures 1.1 - 1.4 show four decentralized detection configurations that are treated in this dissertation. The parallel fusion system shown in Fig. 1.1 represents the simplest type of decentralized detection structure. It consists of n local detectors and a fusion center. Due to the simplicity of its structure, the design of the system subject to different design criteria has received most of the interest in the literature. This system is used for hypothesis testing. In this dissertation, we restrict our attention to binary hypothesis testing problems where the null hypothesis H_0 is tested against the alternative hypothesis H_1. The system receives n observations X_1, X_2,..., X_n in which X_i is the observation received by the local detector LD_i (in the classical detection theory, it is assumed that all these raw observations are available at one central location and that the decision is made based on the entire set of observations). Due to the presence of noise in a practical system, these observations are assumed to be continuous random variables with conditional density functions $p_j(x_i)$; j=0,1 , i=0,1,...,n. From an information theoretic point of view, the transmission of the information in these observations from the local detectors to the global decision maker requires infinite capacity channels. A requirement which is not practically attainable. To overcome this problem, each local detector processes the locally received observation X_i and transmits a compressed version z_i of the data to the global decision maker. The way in

2
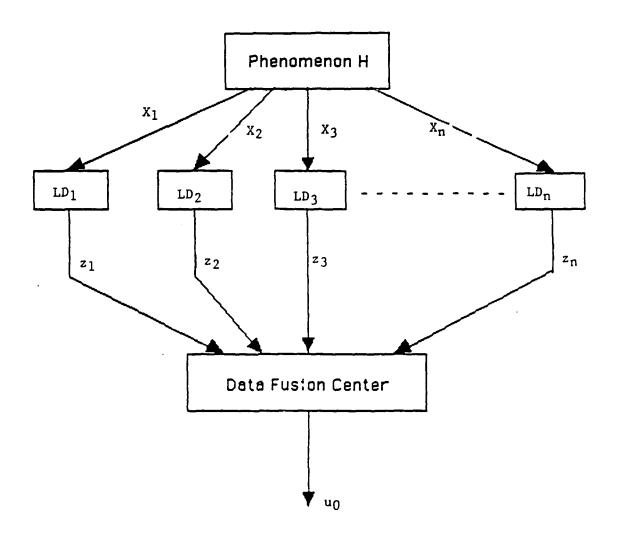
Fig. 1.1. Decentralized detection system with fusion, S_1.
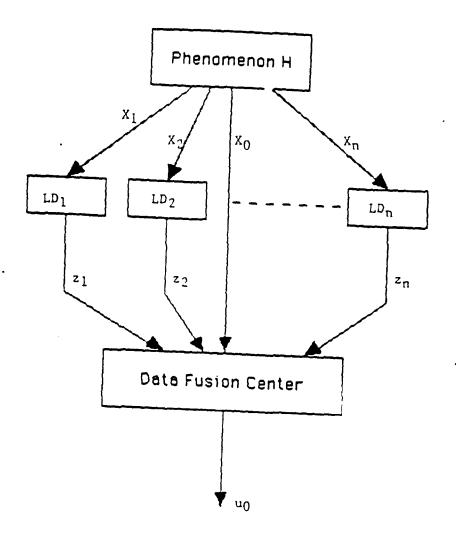
Fig. 1.2. Decentralized detection system with fusion and side information, S_2.
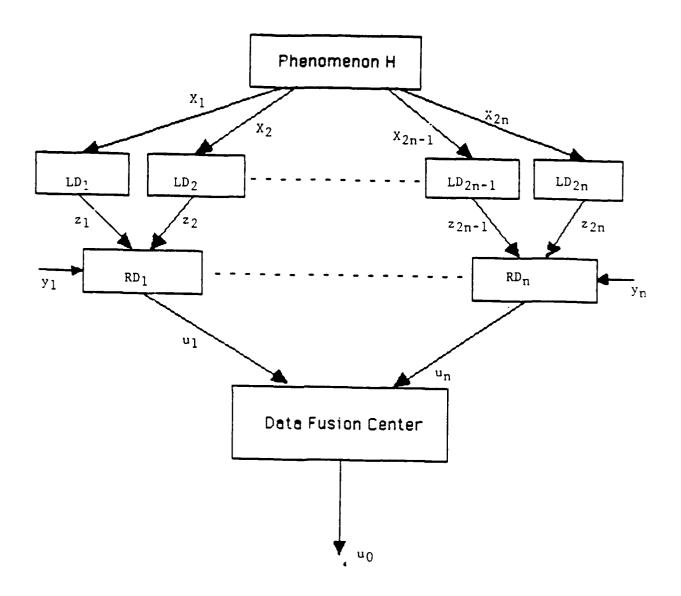
Fig. 1.3. Hierarchical network topology with side information at the regional decision makers, S_3.
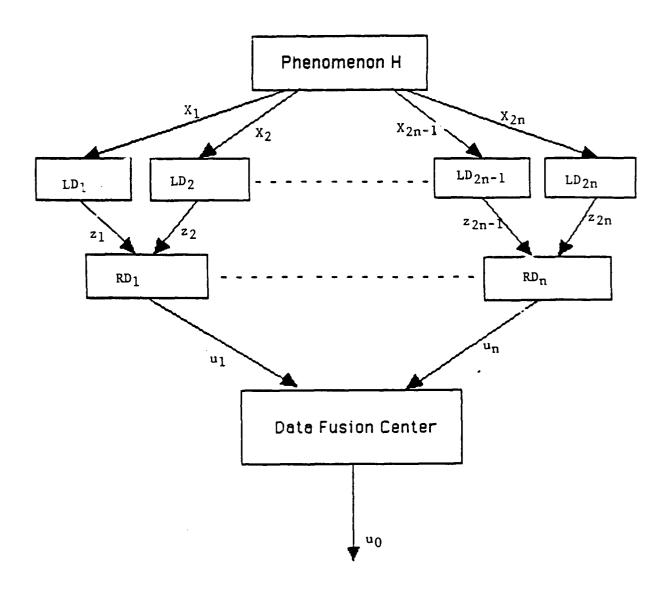
Fig. 1.4.Hierarchical network topology.

which data reduction takes place can be in the form of a hard decision or a soft decision. In the first case, the local decision z_i takes on one of two possible values indicating the presence of hypothesis H_0 or hypothesis H_1 as determined by the local detector LD_i. In the second case, the observation space of each observation X_i is partitioned into M nonoverlapping regions. The local decision z_i, correspondingly, takes on one of M possible values depending upon the region in which the observation X_i falls in. In this case, the local decision z_i does not contain explicit information about the hypothesis present. In both the hard decision and the soft decision cases, the output alphabet is finite and each element in this alphabet has a finite probability of occurrence. This means that the entropy of the alphabet is finite and, therefore, a finite capacity channel can be employed. The local decisions (hard or soft) are sent over the bandlimited channels to the global decision maker, which is also referred to as the fusion center. Based on the decision vector $\underline{U}=[z_1, z_2,...,z_n]$ whose elements are the decisions made by the individual sensors, the fusion center makes the final decision on whether hypothesis H_0 or hypothesis H_1 is true.

Figure 1.2 shows a variation of the parallel fusion system shown in Figure 1.1. Here, the global decision maker receives a local observation of its own (side information) in addition to the local decisions z_i, i=1,2,...,n. The observation vector based on which the fusion center makes the final decision is the augmented vector $[\underline{U}X_0]$ of the local decisions z_i and the observation X_0 at the fusion center. Intuitively, the performance of this system is expected to be better than the performance of the parallel fusion system with n local detectors. Here the information in the observation X_0 is fully utilized by the fusion center and is not compressed by a local detector. In the interesting special case when n =1, the system reduces to a two-stage decentralized serial configuration.

Figure 1.3 shows a hierarchical system with local and regional decision makers. The system consists of 2n local decision makers, n regional decision makers (RD's) and a global decision maker. Local detectors LD_{2i-1} and LD_{2i} process their locally received observations X_{2i-1} and X_{2i} and forward their decisions z_{2i-1} and z_{2i} to an intermediate regional decision maker RD_i , i=1,...,n. The regional detector RD_i combines the two local decisions along with its directly received observation Y_i to make the regional decision u_i. The

decision vector $\underline{U}=[u_1...u_n]$ is used by the global decision maker to make the final decision u_0. The case of more than two local detectors per regional decision maker can also be considered but is not treated in this dissertation. Figure 1.4 shows another hierarchical decentralized system. The difference between this system and the one shown in Figure 1.3 is that the regional decision makers receive no observations of their own. Therefore, they have to make their decisions solely on the basis of the local decisions they receive. The local decisions are made on the basis of the local observations and the global decision is made on the basis of the decisions received from the regional decision makers. However, due to the unavailability of observations at the regional detectors, the performance of this system is expected to be inferior to the performance of the hierarchical system with side information at the regional level.

1.2 Literature Survey

In this section, we briefly review related work on some topics that are treated in this dissertation.

1.2.1 Decentralized Detection Systems

The design of decentralized detection systems has been dealt with by a number of authors, e.g. [4-29]. The parallel fusion system, in particular, has gained most of the interest in the literature. Tenney and Sandell [5] addressed the problem of structuring a set of decision makers and communication links in a manner which leads to effective management of a complex, large scale system in real time. In a follow up paper [6], they developed mechanisms based on the interactions between subsystems to coordinate the making of decisions which are "best" in some system-wide sense. One interesting application of their work is the area of distributed detection systems. Tenney and Sandell [7] treated the two-sensor decentralized detection problem with no data fusion from a Bayesian point of view. Costs were assigned to reflect the course of action of each local detector. The local decision rules were chosen such that the average cost is minimized. The local decision rules were shown to be likelihood ratio tests of the sensor observations for conditionally independent observations. Sadjadi [8] extended the theory in [7] to encompass the M hy-

8

pothesis testing problem with n local detectors. Here again, the fusion center was not a part of the optimization. Assuming known sensor thresholds, Chair and Varshney [9] developed a minimum average cost algorithm for combining the sensor decisions in an n-sensor system at the fusion center. Optimization of the entire system was considered by Hoballah and Varshney [10] where they obtained a person-by-person optimal solution to the parallel network with n local detectors and a fusion center. The person-by-person optimal design for the binary hypothesis testing problem requires the joint solution of (2^n+n) simultaneous nonlinear equations. An iterative algorithm for the solution of the person-by-person optimality equations was proposed in [11]. This algorithm is based on the Gauss-Siedel method for the solution of coupled nonlinear equations. The optimization of the entire system was also considered by Reibman and Nolte [12] where an exhaustive search is performed over all the fusion rules in order to determine the overall minimum cost solution. In the exhaustive search method, the fusion center is fixed at a particular fusion rule and a set of n coupled nonlinear equations are solved to determine the n local thresholds. This process has to be repeated for all the permissible fusion rules. The fusion rule along with the n local thresholds that yield the smallest possible cost are the optimum system design parameters. The number of fusion rules grows quite rapidly with n. Thomopoulos, Viswanathan, and Bougoulias [13] showed that the number of fusion rules to be examined for various values of n is given as shown in the following table.

Number of local detectors	Number of permissible fusion rules
2	2
3	9
4	114
5	6894

Table 1.1

In both the exhaustive method and the person-by-person optimization approach, the

9

amount of computation required grows exponentially. This makes the above design proce-dures computationally difficult and, therefore, the study of decentralized detection systems has been limited to small networks and very few topologies. In this dissertation, we present a computationally simpler approach for the design of decentralized Bayesian de-tection systems. Using this approach, we show that the design of the optimum decentral-ized parallel fusion system reduces to the optimization of a single function of n variables. Chair and Varshney [14] considered the problem of distributed Bayesian hypothesis test-ing with distributed data fusion in which data fusion is performed at each site.

The Neyman -Pearson criterion has also been used in the design of decentralized de-tection systems where neither the costs nor the a priori probabilities need to be available to the designer. Srinivasan [15] used the Neyman-Pearson criterion to obtain the local deci-sion rules in the parallel fusion network assuming that the fusion center is a combinational logic circuit. Hoballah and Varshney [16] treated the problem in two respects. First, when the fusion rule is known and the objective is to find the local decision rules. Second, when the decision rules at the detectors are given, and the objective is to find the optimum fusion rule. Thomopoulos, Viswanathan, and Bougoulias [17] employed the Neyman-Pearson criterion where both the decision made by each individual sensor and the global decision made by the fusion center are based on the Neyman-Pearson test. Tsitsiklis [18] addressed the question of concavity of the receiver operating characteristic of the system. He found that for a given strategy, the receiver operating characteristic is not necessarily a concave function (a numerical example is provided in [19]). However, concavity can be achieved by randomizing with respect to the possible strategies. A similar result was obtained by Willet and Warren [20]. Viswanathan, Thomopoulos, and Tumuluri [21] applied the Ney-man-Pearson criterion to the design of the serial decentralized configuration. They found that for the case of two sensors, the optimal serial network has a better performance than the parallel scheme (better here refers to higher probability of detection for the same false alarm probability). While this interesting result is true for the case of two sensors, the nu-merical examples provided by [21] show that this result is not true, in general, for systems with more than two sensors.

The Bayesian formulation has also been applied to the design of decentralized detection structures other than the parallel fusion network. Ekchian and Tenney [22] derived the necessary conditions for the thresholds to satisfy in order to minimize a given cost function for a number of configurations including the tandem and the tree-hierarchical topologies. No numerical results were provided. Reibman and Nolte [23] applied the exhaustive search method to design specific decentralized detection configurations. Tsitsiklis [24] and Varshney [25] provided an overview of the recent advances in the theory of decentralized detection systems.

1.2.2 The Class of Ali-Silvey Distance Measures and the Detection Problem

The class of Ali-Silvey distance measures has played an important role in the design of quantizers for hypothesis testing. Recently, it has also been employed in the design of decentralized detection systems. In this dissertation, we apply members of the class of Ali-Silvey distance measures to design suboptimum decentralized detection systems, and compare their performance to that of the optimum decentralized detection system. For the sake of completeness we now define the class of Ali-Silvey distance measures and provide examples of members of this class.

Let $p_0(x)$ and $p_1(x)$ be the conditional probability density functions of the random variable X under hypotheses H_0 and H_1 respectively. A measure of dissimilarity between these density functions can be expressed in terms of the general class of Ali-Silvey distance measures (or the class of f-divergences) which is defined as [30]

$$D(p_0(x), p_1(x)) = f\{E_0[g(l)]\} \qquad (1.1)$$

where

E_0 is the expectation under the hypothesis H_0

g is a convex real-valued function defined on $(0, \infty)$

11

f is an increasing function

$l = p_1(x) / p_0(x)$ is the likelihood ratio function

Examples of measures from this class include

1) The Bhattacharyya distance

$$D_B = -\ln \int_x \sqrt{p_0(x) p_1(x)} \, dx \qquad (1.2)$$

2) The discrimination distance measure

$$D_D = \int_x p_0(x) \ln \frac{p_0(x)}{p_1(x)} \, dx \qquad (1.3)$$

3) The Kolmogorov variational distance

$$D_K = \int_x |p_1(x) - p_0(x)| \, dx \qquad (1.4)$$

4) The J-divergence

$$D_J = \int_x [p_1(x) - p_0(x)] \ln \frac{p_1(x)}{p_0(x)} \, dx \qquad (1.5)$$

The class of Ali-Silvey distance measures has received an increasing interest in the design of quantizers for hypothesis testing. This is due to the strong link between these distance measures and the probability of error (POE). Let $\pi = \{\pi_0, \pi_1\}$ be the set of all permissible pairs of the prior probabilities in a binary hypothesis testing problem. Then, there exists a subset of π for which if the distance between a given set of conditional densities is larger than the distance between another set of conditional densities, then the POE corresponding to the first set is less than the POE corresponding to the second set.

12

This result is known as the Blackwell theorem [31]. Poor and Thomas [32] applied the general class of Ali-Silvey distance measures to the design of a generalized quantizer for binary decision systems. Poor [33] also used these measures as criteria for analyzing the effects of fine data quantization on inferential procedures and for designing quantizers to minimize these detrimental effects. Benitz and Bucklew [34] applied the Chernoff theorem to the design of quantizers that asymptotically minimize the probability of error (The Chernoff theorem differs from the traditional Chernoff bound in that the theorem clearly exhibits the exponential dependence of the bound). A number of authors have applied members of the class of Ali-Silvey distance measures in the design of decentralized detection systems. For example, Longo, Lookabaugh, and Gray [35] have employed the Bhattacharyya distance, whereas Lee and Chao [36] have used the J-divergence to subpartition the decision space when a quality bit is transmitted along with the decision to the fusion center.

1.2.3 Information Theoretic Measures for Quantization and Detection

The quantization of a random variable X for minimum distortion has been studied extensively in the literature (see [39] for a survey of results). When the entropy of the quantizer output is restricted not to exceed a given prescribed value, the problem becomes that of quantization under entropy constraint. The quantization of the random variable X for minimum distortion under entropy constraint was considered by Noll and Zelinski [40] and Farvardin and Modestino [41]. It is known that in detection and estimation problems, quantization for minimum distortion is not the appropriate criterion to use for designing an optimum inferential system [42,43]. The problem of adjusting the threshold in a simple binary hypothesis testing problem under the condition of maximizing the mutual information between the decision and the state of nature was considered by Middleton [44] and Gabriele [45]. Martinez [46] and Hoballah and Varshney [47] have shown that the problem of maximizing the mutual information between the decision and the state of nature is equivalent to applying the Neyman-Pearson criterion for signal detection. As a result, the maximum mutual information detector is a likelihood ratio detector. Hoballah and Varsh-

ney [47] have made use of this result to design a maximum mutual information decentralized detector under the assumption that each local detector makes a single hard decision.

1.2.4 Bounds on the Probability of Error of Optimum Receivers

In hypothesis testing, the performance of optimum receivers is usually expressed in terms of the probability of error. It is well known that the optimum receiver which minimizes the probability of error is the maximum a posteriori probability (MAP) receiver [3]. When an observation x is received, the MAP receiver computes the a posteriori probabilities $P(H_0 \mid x)$ and $P(H_1 \mid x)$ of the two hypotheses and chooses the hypothesis with the larger a posteriori probability. The probability of making an error based on the observation x is given as

$$P\left(Error| x\right) \ = \ min\left(P\left(H_0\middle| x\right), P\left(H_1\middle| x\right)\right) \tag{1.6}$$

Analytic evaluation of the probability of error is very difficult in most applications because it involves the evaluation of the discontinuous function min (.). Instead of evaluating the exact probability of error, tight upper and / or lower bounds can often be determined analytically making it possible to compare the performance of optimum receivers based on these bounds.

A number of upper and lower bounds have been proposed in the literature [51-63]. Because of the indirect relationship between the probability of error and the class of Ali-Silvey distance measures discussed above, a number of the bounds available are expressed in terms of these distance measures. The idea behind this lies, of course, in the Blackwell theorem. Boekee and Van der Lubbe [51] provided upper bounds on the probability of error by considering the f-divergence between the conditional densities under the two hypotheses. They have shown that this upper bound includes many well known bounds in terms of other distance measures. The Bhattacharyya bound [31] is the simplest bound to evaluate. Its simplicity and the fact that closed form expressions for the bound exist for many commonly used distributions, made it an attractive tool. The most common applica-

tions are the design of quantizers for hypothesis testing [32,35] and signal selection [31]. The main disadvantage of the Bhattacharyya bound is that it is a loose bound. The Chernoff bound [52,53] provides an upper bound on the probability of error in terms of a scalar s, $0<s<1$. The tightest bound is obtained by optimizing the upper bound with respect to the scalar s. The difficulty in evaluating the Chernoff bound makes it less attractive than the Bhattacharyya bound. In addition, the Chernoff bound which reduces to the Bhattacharyya bound when s=0.5 does not, in general, provide tighter error bounds than the Bhattacharyya bound [54]. A tighter bound on the probability of error than the Bhattacharyya bound is in terms of the equivocation function [55,53]. Devijver [56] introduced another bound in terms of the so called Bayesian distance. This bound is known to be tighter than both the Bhattacharyya bound and the equivocation bound. As applied to decentralized systems, few results are available in the literature. Kazakos [57] employed the concept of distance measures to obtain bounds on the performance of distributed detection systems. Tsitsiklis [58] considered the decentralized problem when the number of sensors tends to infinity. He showed that it is asymptotically optimal for the sensors to use the same decision rule if they are operating under identical circumstances.

1.3 Dissertation Outline

In this dissertation, we consider some design and analysis aspects of a number of decentralized detection structures. Our main focus is on the Bayesian approach to the design of these systems. The design of these systems involves the design of local and global decision rules. The design of optimum decentralized detection systems based on the Bayesian formulation is considered in detail. The performance of the optimum systems is compared to the performance of suboptimum systems designed based on criteria other than the global optimum Bayesian cost. Upper and lower bounds on the minimum probability of error and the minimum achievable cost are derived for the conventional centralized detection system. A new tight upper bound on the minimum probability of error is presented and applied to design a nearly optimum decentralized detection system. The role of randomization of the global decision rule in the Bayesian problem and in the related Neyman-Pearson problem is also discussed.

15

In Chapter 2, we study the problem of optimum receivers from a Beyesian viewpoint. We derive an alternate representation of the minimum achievable cost of an optimum receiver in terms of a modified form of the Kolmogorov variational distance. Using this representation, we show that randomization of the decision rule is not necessary in the case when the observations assume discrete values.

In Chapter 3, we generalize some known upper and lower bounds on the minimum probability of error to the general Bayesian problem ,i.e., we obtain upper and lower bounds on the minimum average cost of optimum Bayesian receivers. These bounds include the Chernoff bound and the Bhattacharyya bound. We derive a new tight upper bound on the minimum probability of error and generalize it to the general Bayesian problem. New tight lower bounds on the probability of error and minimum average cost of an optimum receiver are also obtained.

In Chapter 4, we consider the analysis and the design of the parallel fusion system from a Bayesian point of view assuming identical local detectors. The analysis and the design of the system are based on the representation of the minimum achievable cost of optimum receivers derived in Chapter 2. Both the hard and the soft decision cases are considered. Using this representation we show that the design of the optimum decentralized detection system reduces to the optimization of a single function of a given number of variables that depend upon the number of quantization levels. The design of suboptimum decentralized detection systems based upon members of the class of Ali-Silvey distance measures are also considered. The performance of these systems are compared to the performance of the optimum system. The design of the decentralized detection system based on the new upper bound on the minimum probability of error is also discussed .

In Chapter 5, we consider the design of the four decentralized detection structures shown in Figures 1.1 - 1.4 based on the Bayesian formulation. The local detectors are assumed to yield hard decisions but are not assumed to be identical. The optimum design of each configuration is shown to reduce to the optimization of a single function of a given number of variables depending upon the configuration.

In Chapter 6, we examine the problem of randomization of the decision rule for de-

tection systems designed based on the Neyman-Pearson criterion. When neither the a priori probabilities nor the costs are known, the Neyman-Pearson criterion becomes useful. We show that when the objective is to design a decentralized detection system that maximizes the global probability of detection for a given global false alarm probability, randomization of the decision rule at the fusion center is not necessary and, in fact, if used deteriorates the system performance.

In Chapter 7 we present a summary of the results obtained in this dissertation plus some concluding remarks. We also discuss some of the problems related to the topics discussed in the dissertation that need to be addressed in the future.

CHAPTER 2

MINIMUM AVERAGE COST RECEIVERS

2.1 Introduction

As indicated in Chapter 1, one of the objectives of this dissertation is to consider the design of decentralized Bayesian detection systems. In order to facilitate the design of these systems we first consider centralized Bayesian detection systems, i.e., the systems in which all the raw observations are processed at one central location. A system (centralized or decentralized) that minimizes the Bayesian cost will be referred to as a minimum average cost (MAC) receiver. In this chapter we will derive some important properties of the classical centralized optimum receiver. These are generalizations of the results available in the literature for the minimum probability of error receiver [31,37]. These results will then be applied to the design and performance evaluation of decentralized detection structures in the following chapters.

Let $f_0(x)$ and $f_1(x)$ be the conditional probability density functions of the random observations X under the two hypotheses H_0 and H_1 to be tested at the receiver. Also, let π_0 and π_1 be the corresponding a priori probabilities of H_0 and H_1 respectively. If C_{ij} ; $i, j=0$, 1, denotes the cost of deciding H_i when H_j is true, then the average cost per decision made by the receiver is given by

$$R = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij} P(H_j, H_i) \tag{2.1}$$

where $P(H_j, H_i)$ is the probability of the joint event that H_j is true and H_i is decided. It is well known [1] that the cost given by (2.1) is minimized when the following decision rule is used

$$\Lambda(x) = \frac{f_1(x)}{f_0(x)} \underset{\underset{H_0}{<}}{\overset{\overset{H_1}{>}}{}} \frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})} = \eta \qquad (2.2)$$

This decision rule partitions the observation space of X into two optimum decision regions Z_k^*; k = 0, 1, such that when x ϵ Z_k^*, H_k is declared true. The minimum average cost can be determined from (2.1) using the optimum decision regions Z_k^*.

2.2 An Alternate Representation for the Minimum Average Cost

To present our approach for the design of minimum average cost (MAC) receivers, we express the optimum decision rule (2.2) and the corresponding minimum average cost given in (2.1) in an alternate way. We begin by expressing (2.1) as

$$R = \sum_{i=0}^{1}\sum_{j=0}^{1} C_{ij}\pi_j P(H_i|H_j) \qquad (2.3)$$

where $P(H_i | H_j)$ is the conditional probability of deciding H_i when H_j is true. Let Z_k, k = 0, 1, be the decision region corresponding to H_k, such that when the observation x ϵ Z_k, H_k is declared true. The a posteriori probabilities $P(H_i | H_j)$ are given as

$$P(H_i|H_j) = \int_{Z_i} f_j(x)\, dx \qquad (2.4)$$

Substituting (2.4) into (2.3) we get

$$R = \sum_{i=0}^{1}\sum_{j=0}^{1} C_{ij}\pi_j \int_{Z_i} f_j(x)\, dx \qquad (2.5)$$

Using Bayes rule, we have

$$\pi_j f_j(x) = P(H_j|x) f(x) \qquad (2.6)$$

19

where

$$f(x) = \pi_0 f_0(x) + \pi_1 f_1(x) \tag{2.7}$$

is the unconditional density function of X. Using (2.6), we express (2.5) as

$$R = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij} \int_{Z_i} P(H_j|x) f(x) \, dx \cdot \tag{2.8}$$

Since C_{ij} is a constant, we interchange the order of integration and the inner summation as

$$R = \sum_{i=0}^{1} \int_{Z_i} \sum_{j=0}^{1} C_{ij} P(H_j|x) f(x) \, dx \tag{2.9}$$

Expanding (2.9) over the outer summation, we get

$$R = \int_{Z_0} \beta_0(x) f(x) \, dx + \int_{Z_1} \beta_1(x) f(x) \, dx \tag{2.10}$$

where

$$\beta_0(x) = C_{00} P(H_0|x) + C_{01} P(H_1|x) \tag{2.11}$$

$$\beta_1(x) = C_{10} P(H_0|x) + C_{11} P(H_1|x) \tag{2.12}$$

The result in (2.10) reveals that the average cost given by (2.1) can, equivalently, be represented as the expected value of the random variable $\beta(x)$ defined as

$$\beta(x) = \begin{cases} \beta_0(x) & \text{if } x \in Z_0 \\ \beta_1(x) & \text{if } x \in Z_1 \end{cases} \qquad (2.13)$$

The quantities $\beta_0(x)$ and $\beta_1(x)$ represent the conditional costs assigned to each point x in the observation space. When an observation x is received, the receiver computes the conditional costs $\beta_0(x)$ and $\beta_1(x)$. If the objective at the receiver is to minimize the cost, then it should select the hypothesis with the smaller conditional cost. The optimum decision rule (2.2) based on the conditional cost formulation becomes

$$\beta_0(x) \overset{\overset{H_1}{>}}{\underset{\underset{H_0}{<}}{}} \beta_1(x) \qquad (2.14)$$

Denote by r(x) the conditional cost of the MAC receiver based on an observation x. Then r(x) can be expressed as

$$r(x) = min(\beta_0(x), \beta_1(x)) \qquad (2.15)$$

Using the mathematical identity

$$min(a, b) = \frac{1}{2}(a+b) - \frac{1}{2}|a-b| \qquad (2.16)$$

we can rewrite (2.15) as

$$r(x) = \frac{1}{2}[\beta_0(x) + \beta_1(x)] - \frac{1}{2}|\beta_0(x) - \beta_1(x)| \qquad (2.17)$$

Substituting (2.11) and (2.12) into (2.17) we get

$$r(x) = \frac{1}{2}[C_{00}P(H_0|x) + C_{10}P(H_0|x) + C_{01}P(H_1|x) + C_{11}P(H_1|x)]$$

$$-\frac{1}{2}|C_{00}P(H_0|x) - C_{10}P(H_0|x) + C_{01}P(H_1|x) - C_{11}P(H_1|x)| \qquad (2.18)$$

The minimum average cost R_m is the expected value of the conditional cost $r(x)$, that is

$$R_m = \int_x r(x)[\pi_0 f_0(x) + \pi_1 f_1(x)]dx \qquad (2.19)$$

For notational simplicity, the subscript m will be omitted from R_m in the rest of this dissertation. Using (2.6), we can evaluate the a posteriori probabilities as

$$P(H_j|x) = \frac{\pi_j f_j(x)}{f(x)} \qquad (2.20)$$

Substituting (2.20) into (2.18) we obtain

$$r(x) = \frac{1}{2f(x)}[(C_{00} + C_{10})\pi_0 f_0(x) + (C_{01} + C_{11})\pi_1 f_1(x)]$$

$$-\frac{1}{2f(x)}|(C_{00} - C_{10})\pi_0 f_0(x) + (C_{01} - C_{11})\pi_1 f_1(x)| \qquad (2.21)$$

Substituting (2.21) into (2.19) and integrating term by term we obtain the following expression for the minimum average cost R

$$R = R_0 - \frac{1}{2}\int_x |C_1 f_1(x) - C_0 f_0(x)|dx \qquad (2.22)$$

22

where

$$R_0 = \frac{1}{2}(C_{00} + C_{10})\pi_0 + \frac{1}{2}(C_{01} + C_{11})\pi_1$$

$$C_1 = (C_{01} - C_{11})\pi_1$$

$$C_0 = (C_{10} - C_{00})\pi_0$$

When the random variable X assumes discrete values, the MAC , R, becomes

$$R = R_0 - \frac{1}{2}\sum_{all\ i}|C_1 P(X = x_i | H_1) - C_0 P(X = x_i | H_0)| \qquad (2.23)$$

2.3 Randomization in Minimum Average Cost Receivers

When the random variable X assumes continuous values, the contribution to the minimum average cost of those values of X that satisfy equality in the decision rule (2.14) is zero. This is because these values occur with zero probability. Therefore, they can be assigned to either one of the decision regions without affecting the minimum average cost. When X is discrete, the values of X that satisfy equality in (2.14) occur with finite probability. Therefore, randomization of the decision rule may become necessary. While partitioning the observation space of the discrete random variable X into the decision regions corresponding to H_0 and H_1, the receiver uses the decision rule (2.2) with

$$\Lambda(x_i) = \frac{P(X = x_i | H_1)}{P(X = x_i | H_0)} \qquad (2.24)$$

Let $K = \{k_1, k_2, ..., k_l\}$ be the set of those values of X that satisfy the relation

23

$$\frac{P(X=k_i|H_1)}{P(X=k_i|H_0)} = \frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})} \tag{2.25}$$

The complement set K^c represents all the remaining values that the random variable X assumes. Using Bayes rule we can express (2.25) as

$$\frac{P(H_1|X=k_i)}{P(H_0|X=k_i)} = \frac{(C_{10}-C_{00})}{(C_{01}-C_{11})} \tag{2.26}$$

The conditional costs $\beta_0(x)$ and $\beta_1(x)$ for the outcomes that satisfy (2.26) are given by

$$\beta_0(k_i) = C_{00}P(H_0|X=k_i) + C_{01}P(H_1|X=k_i) \tag{2.27}$$

$$\beta_1(k_i) = C_{10}P(H_0|X=k_i) + C_{11}P(H_1|X=k_i) \tag{2.28}$$

Using (2.26) we get

$$P(H_1|X=k_i) = P(H_0|X=k_i)\frac{(C_{10}-C_{00})}{(C_{01}-C_{11})} \tag{2.29}$$

Substituting (2.29) into (2.27) and (2.28), and simplifying we get

$$\beta_0(k_i) = \beta_1(k_i) = P(H_0|X=k_i)\frac{(C_{10}C_{01}-C_{11}C_{00})}{C_{01}-C_{11}}, k_i \in K \tag{2.30}$$

The result in (2.30) indicates that such boundary outcomes are equally costly and can be assigned arbitrarily to any one of the decision regions. Using (2.23) we will show that the contribution of these outcomes to the MAC is zero. This can be done by expanding (2.23) as

$$R = R_0 - \frac{1}{2} \sum_{i=1}^{l} |C_1 P (X= k_i| H_1) - C_0 P (X= k_i| H_0)|$$

$$- \frac{1}{2} \sum_{x_i \in K^c} |C_1 P (X= x_i| H_1) - C_0 P (X= x_i| H_0)| \qquad (2.31)$$

In (2.31), the second term on the right hand side is the sum over all those values of X that are elements of K, and the third term is the sum over all the values that are elements of K^c. Using (2.25) we get

$$\sum_{i=1}^{l} |C_1 P (X= k_i| H_1) - C_0 P (X= k_i| H_0)| = 0 \qquad (2.32)$$

The result in (2.32) shows that it is, in fact, immaterial as to whether the outcomes that satisfy (2.25) are assigned to the decision space of H_0 or to the decision space of H_1 since their contribution to the MAC is zero. Thus, randomization is not necessary. Therefore, whether the random variable X is continuous or discrete, the minimum average cost is not affected by those values of X that satisfy equality in the likelihood ratio test.

2.4 Minimum Probability of Error Receivers

As an important application of the general Bayesian formulation developed in Section 2.2, we consider the special cost assignment $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$. This corresponds to the minimum probability of error criterion which is widely used. In this case, the conditional costs $\beta_0(x)$ and $\beta_1(x)$ given by (2.11) and (2.12) become

$$\beta_0 (x) = P (H_1| x) \qquad (2.33)$$

25

$$\beta_1(x) = P(H_0 | x) \tag{2.34}$$

and the decision rule (2.14) becomes

$$P(H_1 | x) \underset{\underset{H_0}{<}}{\overset{\overset{H_1}{>}}{}} P(H_0 | x) \tag{2.35}$$

Decision rule (2.35) indicates that the MAC receiver reduces to a maximum a posteriori probability (MAP) receiver in this special case. When an observation x is received, the receiver computes the a posteriori probabilities $P(H_0|x)$ and $P(H_1|x)$ and chooses the hypothesis with the larger a posteriori probability. The function r(x) defined in (2.15) representing the conditional cost of the MAC receiver based on an observation x becomes

$$r(x) = min(P(H_0 | x), P(H_1 | x)) \tag{2.36}$$

This function now defines the probability that the MAP receiver makes an error based on an observation x. The minimum probability of error (MPOE) is obtained from (2.22) and is given by

$$P(E) = \frac{1}{2} - \frac{1}{2} \int_x |\pi_1 f_1(x) - \pi_0 f_0(x)| dx \tag{2.37}$$

which is the same expression as obtained by Kailath [31] and Toussaint [37] using a different approach.

In this chapter, we have obtained an expression for the minimum average cost for an optimum receiver. This result is a generalization of a similar result for the minimum probability of error based on the Kolmogorov variational distance. The result obtained here will be employed in the next few chapters for decentralized detection systems.

CHAPTER 3

BOUNDS ON THE PERFORMANCE OF OPTIMUM RECEIVERS

3.1 Introduction

In hypothesis testing, the performance of optimum receivers is usually expressed in terms of the probability of error. In this chapter, we concentrate on some performance aspects of optimum receivers. We consider the binary hypothesis testing problem in which hypothesis H_0 with a priori probability π_0 is tested against hypothesis H_1 with a priori probability π_1. The decision is made based on a random observation X with conditional probability density functions $f_0(x)$ and $f_1(x)$ when H_0 and H_1 are true respectively. It is well known that the optimum receiver which minimizes the probability of error is the maximum a posteriori probability (MAP) receiver [3]. When an observation x is received,the MAP receiver computes the a posteriori probabilities $P(H_0|x)$ and $P(H_1|x)$ and chooses the hypothesis with the larger a posteriori probability. In Chapter 2, we derived closed form expressions for the minimum probability of error (MPOE) and the general minimum average cost (MAC) of the optimum receiver. Analytic evaluation of these expressions is very difficult in most cases because it involves the evaluation of the discontinuous function min(.). Instead of evaluating the exact minimum probability of error and the exact minimum average cost, tight upper and / or lower bounds can often be determined analytically in an easier fashion making it possible to compare the performance of optimum receivers based on these bounds.

A number of upper and lower bounds on the MPOE of optimum receivers have been proposed in the literature [51-63]. Because of the indirect relationship between the probability of error and the class of Ali-Silvey distance measures (or the f-divergence) [31], a number of the bounds available are expressed in terms of these distance measures. The idea behind this relation lies in the result known as the Blackwell theorem.This theorem

states the following. Let π be the set of all permissible pairs of the prior probabilities π_0 and π_1 in a binary hypothesis testing problem. Then there exists a subset of π for which if the distance between a given set of conditional density functions is larger than the distance between another set of conditional density functions, then the probability of error corresponding to the first set is less than the probability of error corresponding to the second set. Boekee and Van der Lubbe [51] provided upper bounds on the MPOE of optimum receivers by considering the f-divergence between the conditional densities under the two hypotheses. They have shown that this upper bound includes many well known bounds in terms of other distance measures.

As discussed before, almost all of the literature in this area has been limited to finding bounds on the MPOE, and little can be found on the general Bayesian problem in which arbitrary costs are assigned to each course of action in the decision process. In this chapter we extend some of the known bounds on the MPOE to the general Bayesian problem. In addition, we obtain a new upper bound on the MPOE which is tighter than the previously available bounds. We also obtain a tight lower bound on MPOE. In Section 3.2, we derive an upper bound on the MAC which is a generalization of the Chernoff bound on the MPOE. In Section 3.3, we derive simple upper and lower bounds on the MAC in terms of the Bhattacharyya coefficient. In Section 3.4, we introduce our new upper bound on the MPOE. This bound is shown to be tighter than the available bounds such as the Bhattacharyya bound and the equivocation bound. In Section 3.5, we extend this bound to the restricted Bayesian problem in which $C_{00} = C_{11}$ and $C_{10} = C_{01}$. In Section 3.6, we use the upper bound derived in Section 3.4 to obtain a new tight lower bound on the MPOE. In Section 3.7, we present a numerical example where we compare the exact minimum probability of error with the new upper and lower bounds on the minimum probability of error. Section 3.8 contains a summary of the results obtained in this chapter.

3.2 Generalized Chernoff Bound

In this section we derive a new upper bound on the system MAC which is a generalization of the Chernoff bound on the MPOE available in the literature [52,53]. We begin

with (2.11) and (2.12) which are repeated here for convenience

$$\beta_0(x) = C_{00} P(H_0|x) + C_{01} P(H_1|x) \tag{3.1}$$

$$\beta_1(x) = C_{10} P(H_0|x) + C_{11} P(H_1|x) \tag{3.2}$$

Using Bayes rule we can express the a posteriori probabilities $P(H_0|x)$ and $P(H_1|x)$ as

$$P(H_0|x) = \frac{\pi_0 f_0(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)} \tag{3.3}$$

$$P(H_1|x) = \frac{\pi_1 f_1(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)} \tag{3.4}$$

Using (3.3) and (3.4), the expressions for $\beta_0(x)$ and $\beta_1(x)$ in (3.1) and (3.2) become

$$\beta_0(x) = \frac{C_{00} \pi_0 f_0(x) + C_{01} \pi_1 f_1(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)} \tag{3.5}$$

$$\beta_1(x) = \frac{C_{10} \pi_0 f_0(x) + C_{11} \pi_1 f_1(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)} \tag{3.6}$$

The MAC receiver computes the quantities $\beta_0(x)$ and $\beta_1(x)$ and makes the decision according to decision rule (2.14). The conditional cost based on an observation x is given by

$$r(x) = min(\beta_0(x), \beta_1(x)) \tag{3.7}$$

For any two positive real numbers a and b, we have the following inequality

29

$$min\,(a, b) \le a^s b^{1-s}, 0 \le s \le 1 \qquad (3.8)$$

Making use of (3.8) we can obtain an upper bound on the conditional cost r(x) given in (3.7) as

$$r(x) \le (\beta_0(x))^s (\beta_1(x))^{1-s}, 0 \le s \le 1 \qquad (3.10)$$

The unconditional cost, R, is given by

$$R = \int_x min\,(\beta_0(x), \beta_1(x))\,(\pi_0 f_0(x) + \pi_1 f_1(x))\,dx \qquad (3.11)$$

Therefore, an upper bound on R is

$$R \le \int_x (\beta_0(x))^s (\beta_1(x))^{1-s} (\pi_0 f_0(x) + \pi_1 f_1(x))\,dx, 0 \le s \le 1. \qquad (3.12)$$

Substituting (3.5) and (3.6) in (3.12) we get

$$R \le \int_x (C_{00}\pi_0 f_0(x) + C_{01}\pi_1 f_1(x))^s (C_{10}\pi_0 f_0(x) + C_{11}\pi_1 f_1(x))^{1-s}dx, 0 \le s \le 1.$$
$$(3.13)$$

The upper bound in (3.13) is true for any value of s in the range $0 \le s \le 1$. The tightest bound is obtained by finding the particular value of s which minimizes the right hand side of (3.13). Therefore, the tightest upper bound of this form on MAC is given as

$$R \le \min_{0 \le s \le 1} \int_x (C_{00}\pi_0 f_0(x) + C_{01}\pi_1 f_1(x))^s (C_{10}\pi_0 f_0(x) + C_{11}\pi_1 f_1(x))^{1-s}dx$$
$$(3.14)$$

This is a generalization of the well known Chernoff bound on the minimum probability of error. Using the special cost assignments $C_{00} = C_{11} = 0$ and $C_{10} = C_{01} = 1$, we can obtain it from (3.14) as

$$P(E) \leq \min_{0 \leq s \leq 1} \int_x (\pi_1 f_1(x))^s (\pi_0 f_0(x))^{1-s} dx \qquad (3.15)$$

Let s* be the specific value of s that achieves the minimum in (3.15). It has been shown in [53] that (3.15) can be written as

$$P(E) \leq \pi_1^{s^*} \pi_0^{1-s^*} \rho^* \qquad (3.16)$$

where

$$\rho^* = \min_{0 \leq s \leq 1} \int_x (f_1(x))^s (f_0(x))^{1-s} dx \qquad (3.17)$$

is the Chernoff coefficient. Equation (3.16) represents the usual form of the Chernoff bound on the minimum probability of error.

3.3 Upper and Lower Bounds on the Minimum Average Cost in Terms of the Bhattacharyya Coefficient

The evaluation of the generalized Chernoff bound in (3.14) is a difficult task. Therefore it is worthwhile to derive upper and lower bounds on the MAC that are simpler to evaluate. Here we derive bounds in terms of the Bhattacharyya coefficient. This coefficient is obtained from (3.17) by using the specific value of s = 0.5, i.e., the Bhattacharyya coefficient is given as

$$\rho_B = \int_x \sqrt{f_1(x) f_0(x)} \, dx \qquad (3.18)$$

31

The Bhattacharyya coefficient is relatively simple to evaluate and has closed form expressions for many commonly used distributions. These advantages made the Bhattacharyya bound on the minimum probability of error an attractive tool in the design of communication systems. The most common applications are the design of quantizers for hypothesis testing [32] and signal selection [31]. The main disadvantage of the Bhattacharyya bound is that it is a loose bound. In this section we derive upper and lower bounds on the MAC based on the Bhattacharyya coefficient which are generalizations of the available bounds on the MPOE. These generalized bounds are based upon the following representation of the MAC derived in Chapter 2

$$R = R_0 - \frac{1}{2}\int_x |C_1 f_1(x) - C_0 f_0(x)| \, dx \qquad (3.19)$$

Here we make the usual assumption that an incorrect decision is more costly than a correct decision, i.e., we assume that $C_{10} > C_{00}$ and $C_{01} > C_{11}$. Equation (3.19) can also be written as

$$\int_x |C_1 f_1(x) - C_0 f_0(x)| \, dx = 2R_0 - 2R \qquad (3.20)$$

Next, we obtain the upper and lower bounds on the MAC.

3.3.1 Upper Bound on the MAC

To find an upper bound on the MAC we make use of the following inequality which is true for any two positive real numbers a and b

$$|a - b| \geq [\sqrt{a} - \sqrt{b}]^2 \qquad (3.21)$$

With the aid of (3.21),the left hand side of (3.20) can be bounded as

$$\int_x |C_1 f_1(x) - C_0 f_0(x)| \, dx \geq \int_x [\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}]^2 \, dx \qquad (3.22)$$

Expanding the right hand side of (3.22) we get

$$\int_x [\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}]^2 \, dx = \int_x (C_1 f_1(x) + C_0 f_0(x) - 2\sqrt{C_1 C_0}\sqrt{f_1(x) f_0(x)}) \, dx \qquad (3.23)$$

Integrating term by term on the right hand side of (3.23) we get

$$\int_x [\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}]^2 \, dx = C_1 + C_0 - 2\sqrt{C_1 C_0}\rho_B \qquad (3.24)$$

where ρ_B is the Bhattacharyya coefficient defined in (3.18). From (3.20) and (3.22) we get

$$2R_0 - 2R \geq \int_x [\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}]^2 \, dx \qquad (3.25)$$

Substituting (3.24) into (3.25) we obtain

$$2R_0 - 2R \geq C_1 + C_0 - 2\sqrt{C_1 C_0}\rho_B \qquad (3.26)$$

Rearranging the terms in (3.26) we obtain the following upper bound on R

$$R \leq R_0 - \frac{1}{2}(C_1 + C_0) + \sqrt{C_1 C_0}\rho_B \qquad (3.27)$$

Using the definitions of R_0, C_1 and C_0 in (2.2), the upper bound in (3.27) becomes

$$R \leq C_{00}\pi_0 + C_{11}\pi_1 + \sqrt{\pi_0 \pi_1 (C_{01} - C_{11})(C_{10} - C_{00})}\,\rho_B \qquad (3.28)$$

which is the desired upper bound.

3.3.2 Lower Bound on the MAC

In this subsection we derive a lower bound on the MAC in terms of the Bhatta-charyya coefficient. Here we make use of the Schwartz's inequality

$$\int_t f(t)\, g(t)\, dt \le \left[\int_t (f(t))^2 dt \int_t (g(t))^2 dt \right]^{0.5} \tag{3.29}$$

Squaring both sides of (3.29) we get

$$\left[\int_t f(t)\, g(t)\, dt \right]^2 \le \left[\int_t (f(t))^2 dt \right] \left[\int_t (g(t))^2 dt \right] \tag{3.30}$$

For any two real numbers a and b, we have the following equality

$$(a-b)\,(a+b) = a^2 - b^2 \tag{3.31}$$

Let

$$a = \sqrt{C_1 f_1(x)} \tag{3.32}$$

$$b = \sqrt{C_0 f_0(x)} \tag{3.33}$$

Then (3.31) can be written as

$$(\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)})\,(\sqrt{C_1 f_1(x)} + \sqrt{C_0 f_0(x)}) = C_1 f_1(x) - C_0 f_0(x) \tag{3.34}$$

Taking the absolute value of both sides of (3.34) and integrating with respect to x we get

$$\int_x \left|\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}\right| \left|\sqrt{C_1 f_1(x)} + \sqrt{C_0 f_0(x)}\right| dx = \int_x \left|C_1 f_1(x) - C_0 f_0(x)\right| dx$$

$$(3.35)$$

Squaring both sides of (3.35) we get

$$\left[\int_x \left|\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}\right| \left|\sqrt{C_1 f_1(x)} + \sqrt{C_0 f_0(x)}\right| dx\right]^2 = \left[\int_x \left|C_1 f_1(x) - C_0 f_0(x)\right| dx\right]^2$$

$$(3.36)$$

Let us make the following substitutions in equation (3.30)

$$f(t) = \left|\sqrt{C_1 f_1(t)} - \sqrt{C_0 f_0(t)}\right| \tag{3.37}$$

$$g(t) = \left|\sqrt{C_1 f_1(t)} + \sqrt{C_0 f_0(t)}\right| \tag{3.38}$$

Applying Schwartz's inequality (3.30) to the left hand side of (3.36) with the proper substitutions from (3.37) and (3.38) we get

$$\left[\int_x \left|\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}\right| \left|\sqrt{C_1 f_1(x)} + \sqrt{C_0 f_0(x)}\right| dx\right]^2 \le I_1 I_2 \tag{3.39}$$

where

$$I_1 = \int_x \left|\sqrt{C_1 f_1(x)} - \sqrt{C_0 f_0(x)}\right|^2 dx \tag{3.40}$$

$$I_2 = \int_x \left|\sqrt{C_1 f_1(x)} + \sqrt{C_0 f_0(x)}\right|^2 dx \tag{3.41}$$

35

Expanding the right hand side of (3.40) and (3.41) and integrating we get

$$I_1 = C_1 + C_0 - 2\sqrt{C_1 C_0} \int_x \sqrt{f_1(x) f_0(x)}\, dx \qquad (3.42)$$

$$I_2 = C_1 + C_0 + 2\sqrt{C_1 C_0} \int_x \sqrt{f_1(x) f_0(x)}\, dx \qquad (3.43)$$

It is easily recognized that the integral on the right hand side of (3.42) and (3.43) is the Bhattacharyya coefficient ρ_B. But the left hand side of (3.39) is equal to the right hand side of (3.36). Therefore, we have

$$\left[\int_x |C_1 f_1(x) - C_0 f_0(x)|\, dx \right]^2 \le (C_1 + C_0 - 2\rho_B \sqrt{C_1 C_0})\, (C_1 + C_0 + 2\rho_B \sqrt{C_1 C_0})$$
$$(3.44)$$

Taking the square root of both sides of (3.44) and simplifying we get

$$\int_x |C_1 f_1(x) - C_0 f_0(x)|\, dx \le \sqrt{[C_1 + C_0]^2 - 4 C_1 C_0 \rho_B^2} \qquad (3.45)$$

Using (3.20) we obtain

$$2R_0 - 2R \le \sqrt{[C_1 + C_0]^2 - 4 C_1 C_0 \rho_B^2} \qquad (3.46)$$

Rearranging the terms in (3.46) we obtain the following lower bound on R

$$R \ge R_0 - \frac{1}{2}\sqrt{[C_1 + C_0]^2 - 4 C_1 C_0 \rho_B^2} \qquad (3.47)$$

Substituting for R_0, C_0 and C_1 from (2.22), we have

36

$$R \geq \frac{1}{2} (C_{00} + C_{10}) \pi_0 + \frac{1}{2} (C_{01} + C_{11}) \pi_1$$

$$-\frac{1}{2} \sqrt{[(C_{01} - C_{11}) \pi_1 + (C_{10} - C_{00}) \pi_0]^2 - 4\pi_0\pi_1 (C_{01} - C_{11}) (C_{10} - C_{00}) \rho_B^2}$$

$$(3.48)$$

3.3.3 Minimum Probability of Error Bounds

Here, we consider the minimum probability of error case, i.e., when $C_{00} = C_{11} = 0$, $C_{10} = C_{01} = 1$ and show that the above upper and lower bounds in terms of the Bhattacharyya coefficient reduce to the results available in the literature. In this case, the upper bound in (3.28) simplifies to

$$P(E) \leq \sqrt{\pi_0\pi_1} \rho_B \qquad (3.49)$$

and the lower bound in (3.48) becomes

$$P(E) \geq \frac{1}{2} - \frac{1}{2} \sqrt{1 - 4\pi_0\pi_1\rho_B^2} \qquad (3.50)$$

Equations (3.49) and (3.50) represent the upper and lower bounds on the probability of error which are well known [31]. These bounds are convenient from a computational standpoint but are rather loose. Therefore, we derive tighter upper and lower bounds on the MPOE in the rest of this chapter.

3.4 A Tight Upper Bound on the Probability of Error for Optimum Receivers

For the MAP receiver discussed in Chapter 2, the conditional probability of error based on an observation x is given by

$$P(E|x) = min(P(H_0|x), P(H_1|x)) \tag{3.51}$$

For notational convenience we let

$$p = P(H_0|x), 0 \leq p \leq 1 \tag{3.52}$$

Then we have

$$P(E|x) = min(p, 1-p) \stackrel{\Delta}{=} g(p) \tag{3.53}$$

The function g(p) is plotted in Fig. 3.1 as a function of p. The unconditional probability of error is the expected value of P(E | x) with respect to x, i.e.,

$$P(E) = E_x\{P(E|x)\} = E_x\{g(p)\}$$

$$= \int_x g(p)(\pi_0 f_0(x) + \pi_1 f_1(x)) dx \tag{3.54}$$

The expression for the minimum probability of error given in (3.54) is exact but is computationally undesirable in many applications due to the discontinuity of the function g(p) at p = 0.5. Therefore, we attempt to find a computationally desirable function of p, g*(p), which when substituted for g(p) in (3.54), will provide a close approximation or a tight bound on the probability of error. Let B(g*) denote a bound on the probability of error

$$B(g^*) = \int_x g^*(p)(\pi_0 f_0(x) + \pi_1 f_1(x)) dx \tag{3.55}$$

The closer the function g*(p) is to g(p), the tighter is the bound. In this section, we find an upper bound on the minimum probability of error based on the function g*(p) which satisfies the following conditions

1) g*(p) ≥ g(p) for all values of p in the range 0 ≤ p ≤ 1 such that g*(p = 0) = g*(p = 1) = 0. This condition is consistent with the requirement that g*(p) be an upper bound function.

2) g*(p) should be continuous and differentiable. This condition is desirable to avoid the same computational difficulties as associated with the exact expression based on g(p).

3) g*(p) should be symmetrical about the point p = 0.5. This condition is desirable since g(p) is symmetrical about the point p = 0.5.

4) g*(p = 0.5) = 0.5. This condition is needed since the conditional probability of error P(E|x) cannot be larger than 0.5 and it is equal to 0.5 only when p = 0.5.

5) dg*/dp (p = 0) ≥ 1 and dg*/dp (p = 1) ≤ -1. This condition ensures that the function selected lies above the function g(p) for all values of p in the range 0 ≤ p ≤ 1.

Now we demonstrate that several upper bounds on the minimum probability of error proposed in the literature can be interpreted using the above framework. A similar idea has been pursued by Chen [59]. First, we consider the upper bound based on the Bhattacharyya coefficient and given in (3.49). It can be written as

$$P(E) \leq \int_x \sqrt{\pi_1 f_1(x)\, \pi_0 f_0(x)}\, dx \qquad (3.56)$$

From Bayes rule and the definition of p in (3.52) we have

$$\pi_0 f_0(x) = p\,(\pi_0 f_0(x) + \pi_1 f_1(x)) \qquad (3.57)$$

$$\pi_1 f_1(x) = (1-p)\,(\pi_0 f_0(x) + \pi_1 f_1(x)) \qquad (3.58)$$

39

Substituting (3.57) and (3.58) into (3.56) we get

$$\overline{}) \leq \int_x \sqrt{p(1-p)} \, (\pi_0 f_0(x) + \pi_1 f_1(x)) \, dx$$

$$= \int_x g_B(p) \, (\pi_0 f_0(x) + \pi_1 f_1(x)) \, dx \qquad (3.59)$$

where

$$g_B(p) = \sqrt{p(1-p)} \qquad (3.60)$$

Thus, this upper bound is obtained by replacing g(p) by g_B (p). The function g_B (p) will be referred to as the Bhattacharyya function and is plotted in Fig. 3.1. From Fig. 3.1, we see that g_B(p) is a fairly poor approximation to the function g(p) and, therefore, the Bhatta-charyya bound is a loose one. It can be easily verified that the function g_B (p) satisfies all the five conditions for g*(p) listed above.

Another upper bound on the probability of error that is available in the literature is in terms of the equivocation (an information theoretic measure). This upper bound is given as

$$P(E) \leq -\int_x 0.5 \, [P(H_0|x) \log P(H_0|x) + (1 - P(H_0|x)) \log (1 - P(H_0|x))] f(x) \, dx$$

$$(3.61)$$

In this case, the function g(p) is replaced by the equivocation function g_E(p) given by

$$g_E(p) = -0.5 \, [p \log p + (1-p) \log (1-p)] \qquad (3.62)$$
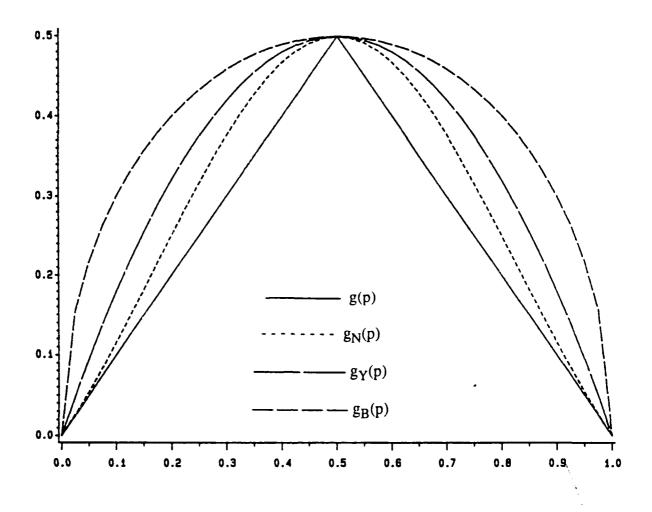
Fig. 3.1 . The exact minimum probability of error, the new upper bound, the sinusoidal, and the Bhattacharyya functions plotted as a function of p.

This function also satisfies the five conditions listed above. It has been shown in [55] that the equivocation bound given by (3.61) gives a tighter bound on the probability of error than the Bhattacharyya bound. Devijver [56] introduced a new bound on the probability of error in terms of what is called the Bayesian distance. This distance is defined as

$$B_Y = E_x \{ \sum_{i=0}^{1} [P(H_i|x)]^2 \} \qquad (3.63)$$

which in our notation reduces to

$$B_Y = E_x \{ p^2 + (1-p)^2 \}$$

$$= E_x \{ 1 - 2p(1-p) \} \qquad (3.64)$$

The probability of error is related to the Bayesian distance through the relation

$$P(E) \le 1 - B_Y$$

$$\le E_x \{ 2p(1-p) \} \qquad (3.65)$$

In this case, the function g(p) is replaced by the function

$$g_Y(p) = 2p(1-p) \qquad (3.66)$$

This function will be called the Bayesian bound function. It has been verified in [56] that the Bayesian bound function yields a tighter bound on the probability of error than the equivocation function. As shown in Fig. 3.1, the function $g_Y(p)$ approximates g(p) much more closely than $g_B(p)$. We have not shown $g_E(p)$ in Fig. 3.1 but it falls between the functions $g_B(p)$ and $g_Y(p)$. Next, we develop a new function that satisfies the desirable conditions indicated earlier and approximates g(p) even more closely. The resulting func-

tion will yield a tighter upper bound on the minimum probability of error.

Consider the sinusoidal function

$$g_s(p) = 0.5 \sin \pi p \qquad (3.67)$$

This function satisfies the desired properties of the approximating function as described earlier. Furthermore, for all values of p in the range $0 \leq p \leq 1$, we have $g_s(p) \leq g_Y(p)$, i.e.,

$$0.5 \sin \pi p \leq 2p(1-p) \qquad (3.68)$$

The result in (3.68) indicates that the sinusoidal function $g_s(p)$ will lead to a tighter upper bound than the Bayesian bound since it is closer to the minimum probability of error function g(p). Differentiating the sinusoidal function in (3.67) and evaluating the derivative at p = 0, we see that the slope at p = 0 is 1.57. This value of the slope is much larger than the slope of g(p) at p = 0 which is equal to 1. We can tighten the bound obtained from (3.67) by weighting the sinusoidal function by a Gaussian function, and we assume that the approximation function is

$$g_N(p) = 0.5 (\sin \pi p) \, exp \, [-\alpha (p - 0.5)^2] \qquad (3.69)$$

The Gaussian function is continuous, differentiable, and symmetrical about the point p = 0.5 as required for the approximating function. The reason for including the Gaussian function in $g_N(p)$ is that it exhibits exponential decay. This property makes it possible to better approximate the exact function g(p) with the sinusoidal and the Gaussian functions combined than with the sinusoidal function alone. The parameter α determines the rate of decay of the Gaussian function. This parameter is to be chosen so that the derivative of $g_N(p)$ at p = 0 is equal to 1. We recall that the derivative of the function g(p) at p = 0 is equal to 1. Therefore, by letting the derivative of $g_N(p)$ at p = 0 to be equal to 1, we are

forcing both functions $g_N(p)$ and $g(p)$ to have the same slope at the points $p = 0$ and $p = 1$. Furthermore, the values of the two functions at the points $p = 0$ and $p = 1$ are equal to zero. Therefore, by having equal values of the two functions as well as equal slopes at the points $p = 0$ and $p = 1$, we can obtain the best approximation possible for the function considered. Differentiating (3.69) with respect to p, setting the derivative at zero equal to one, and solving for α we get $\alpha = 1.8063$. So our new approximating function to obtain the upper bound is

$$g_N(p) = 0.5\,(\sin \pi p)\,exp\,[-1.8063\,(p - 0.5)^2] \qquad (3.70)$$

This function is plotted in Fig. 3.1. In Fig. 3.2, we plot the functions $g(p)$ and $g_s(p)$ versus p. We also plot the function $g_N(p)$ to show the improvement gained by weighting the sinusoidal function $g_s(p)$ by the Gaussian function. By comparing the various approximation functions in Figures 3.1 and 3.2, we see that the new function given in (3.70) approximates the exact function $g(p)$ quite closely, and that it is significantly better than the Bayesian bound function. The probability of error is now upper bounded by

$$P(E) \leq \int_x (0.5\,(\sin \pi p)\,exp\,[-1.8063\,(p - 0.5)^2])\,(\pi_0 f_0(x) + \pi_1 f_1(x))\,dx$$
$$\qquad (3.71)$$

In this section we have presented a new upper bound on the minimum probability of error which is shown to be tighter than the presently available bounds such as the Bhattacharyya, the equivocation, and the Bayesian bounds. Some applications of the new bound in the design of decentralized detection systems will be considered in the next chapter.

3.5 A Tight Upper Bound on the Minimum Average Cost

In the previous section, we derived a tight upper bound on the probability of error for optimum receivers. This is a special case of the general Bayesian problem where the costs
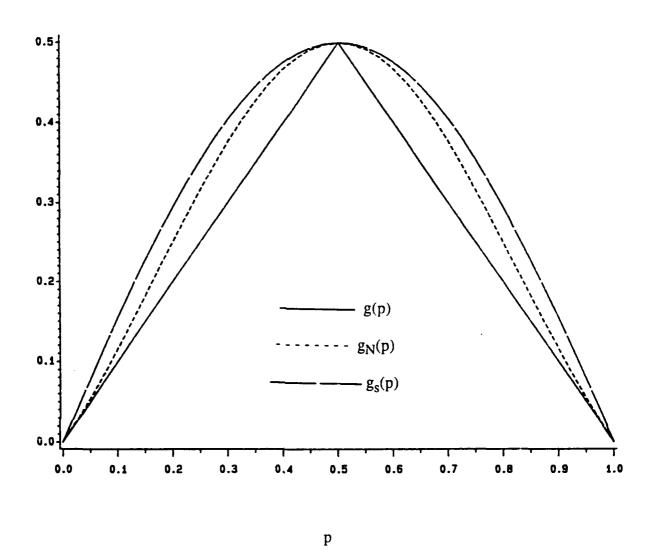
Fig. 3.2 . Sinusoidal and the Gaussian weighted sinusoidal functions.

are assigned as $C_{00} = C_{11} = 0$ and $C_{10} = C_{01} = 1$. In this section we extend the bound to the more general Bayesian hypothesis testing problem with symmetrical costs, i.e., with costs satisfying $C_{00} = C_{11}$ and $C_{10} = C_{01}$. Expressing $\beta_0(x)$ in (3.1) and $\beta_1(x)$ in (3.2) in terms of $P(H_0 \mid x)$, we get

$$\beta_0(x) = C_{00}P(H_0|x) + C_{01}(1 - P(H_0|x)) \tag{3.72}$$

$$\beta_1(x) = C_{10}P(H_0|x) + C_{11}(1 - P(H_0|x)) \tag{3.73}$$

Rearranging terms in (3.72) and (3.73) we get

$$\beta_0(x) - C_{01} = (C_{00} - C_{01})P(H_0|x) \tag{3.74}$$

$$\beta_1(x) - C_{11} = (C_{10} - C_{11})P(H_0|x) \tag{3.75}$$

Dividing (3.74) by (3.75) we get

$$\frac{\beta_0(x) - C_{01}}{\beta_1(x) - C_{11}} = \frac{C_{00} - C_{01}}{C_{10} - C_{11}} \tag{3.76}$$

From (3.76) we express $\beta_0(x)$ in terms of $\beta_1(x)$ as

$$\beta_0(x) = -A\beta_1(x) + B \tag{3.77}$$

where

46

$$A = \frac{C_{01} - C_{00}}{C_{10} - C_{11}} \qquad (3.78)$$

$$B = \frac{C_{01}C_{10} - C_{00}C_{11}}{C_{10} - C_{11}} \qquad (3.79)$$

Using (3.77), the conditional cost in (3.7) becomes

$$r(x) = min(-A\beta_1(x) + B, \beta_1(x)) \qquad (3.80)$$

As $P(H_0 \mid x)$ varies over the range $0 \le P(H_0|x) \le 1$, $\beta_0(x)$ in (3.72) varies over $C_{01} \le \beta_0(x) \le C_{00}$ and $\beta_1(x)$ in (3.73) varies over $C_{11} \le \beta_1(x) \le C_{10}$. For an arbitrary cost assignment the functions $\beta_1(x)$ and $-A\beta_1(x)+B$ as a function of $\beta_1(x)$ are sketched in Fig. 3.3 . Also shown in this figure is the minimum of the two functions which represents r(x). The lack of symmetry of the conditional cost function makes it difficult to extend the upper bound obtained in Section 3.4 to the general Bayesian case. Therefore, we consider the restricted problem with symmetrical costs in which $C_{00} = C_{11}$ and $C_{10} = C_{01}$. In this case A = 1 and $B = C_{01}+C_{11}$. The conditional cost in (3.80) becomes

$$r(x) = min(-\beta_1(x) + C_{01} + C_{11}, \beta_1(x)) \qquad (3.81)$$

The functions $\beta_1(x)$ and $-\beta_1(x)+C_{01}+C_{11}$ are sketched in Fig. 3.4 alongwith the minimum of the two functions. As can be seen from Fig. 3.4, the conditional cost r(x) is a scaled and translated version of the minimum probability of error function g(p). Figure 3.5-a shows the function r(x) - C_{11} resulting from the translation of the function r(x) along the vertical axis. Figure 3.5-b shows the function resulting from translating the function r(x) - C_{11} along the horizontal axis to the origin by C_{11}. The function $g_l(\beta_1)$ sketched in Fig. 3.5-b is represented in terms of $g(\beta_1)$ as

47
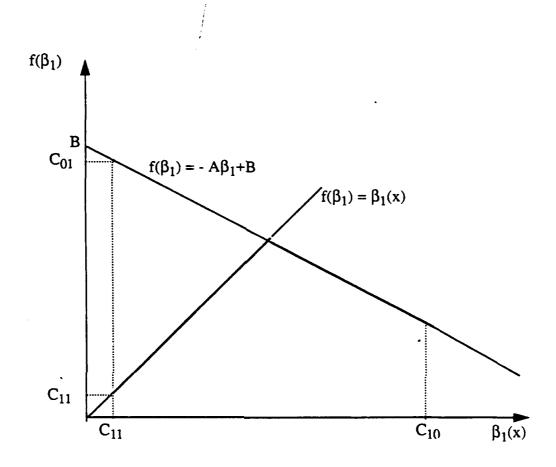
Fig. 3.3. The functions $\beta_1(x)$ and $- \beta_1(x)+C_{10}+C_{11}$ for arbitrary cost assignment.
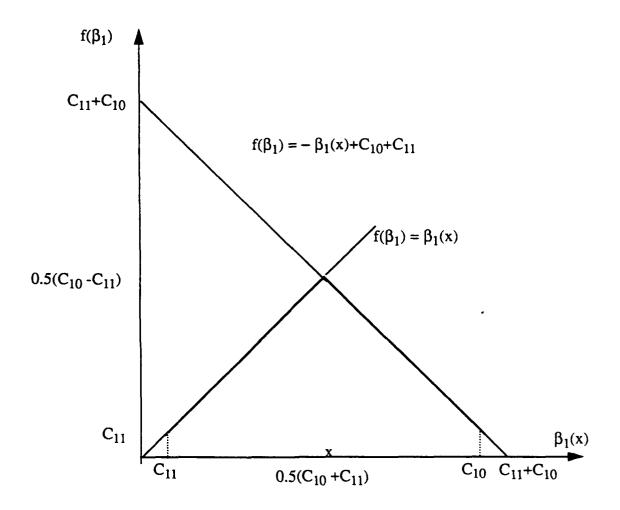
Fig. 3.4. The functions $\beta_1(x)$ and $-\beta_1(x)+C_{10}+C_{11}$ for the case of symmetrical costs.
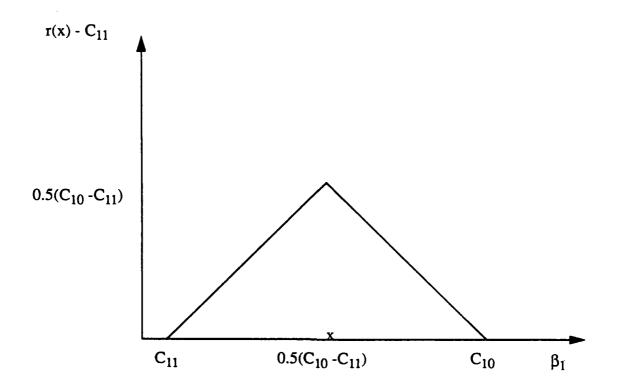
Fig. 3.5-a. Shift the function $r(x)$ in Fig. 3.4 down along the vertical axis by C_{11}.
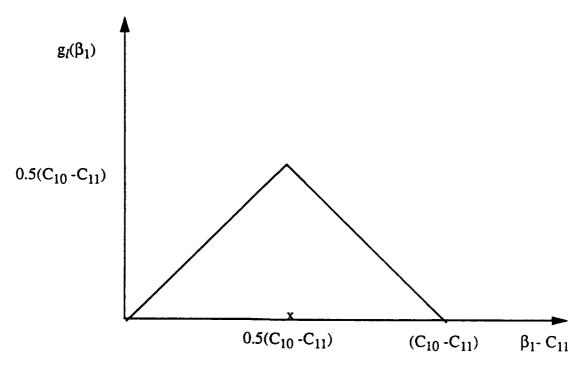
Fig. 3.5-b. Shift the function of Fig. 3.5-a along the horizontal axis to the left by C_{11}.

$$g_l(\beta_1) = (C_{10} - C_{11}) \, g\left(\frac{\beta_1}{C_{10} - C_{11}}\right) \tag{3.82}$$

while the function sketched in Fig. 3.5-a is represented as

$$r(x) - C_{11} = (C_{10} - C_{11}) \, g\left(\frac{\beta_1 - C_{11}}{C_{10} - C_{11}}\right) \tag{3.83}$$

Using (3.83), the function sketched in Fig. 3.4 can be expressed as

$$r(x) = C_{11} + (C_{10} - C_{11}) \, g\left(\frac{\beta_1 - C_{11}}{C_{10} - C_{11}}\right) \tag{3.84}$$

Therefore, using (3.70), we can upper bound the conditional cost r(x) using the function

$$g_R(\beta_1) = C_{11} + (C_{10} - C_{11}) \, g_N\left(\frac{\beta_1 - C_{11}}{C_{10} - C_{11}}\right) \tag{3.85}$$

which can explicitly be expressed as

$$g_R(\beta_1) = C_{11} + 0.5 \, (C_{10} - C_{11}) \sin\frac{\pi \, (\beta_1 - C_{11})}{C_{10} - C_{11}} \, exp\left[-1.8063\left(\frac{\beta_1 - C_{11}}{C_{10} - C_{11}} - 0.5\right)^2\right] \tag{3.86}$$

But β_1 is a function of the random variable X as seen from (3.73). We emphasize this by expressing β_1 as $\beta_1(x)$. Therefore, the MAC is upper bounded by

$$R \le E_x\{g_R(\beta_1(x))\} = \int_x g_R(\beta_1(x)) \, (\pi_0 f_0(x) + \pi_1 f_1(x)) \, dx \tag{3.87}$$

Next, we derive a lower bound on the minimum probability of error.

3.6 A Tight Lower Bound on the Probability of Error for Optimum Receivers

In Section 3.5, we obtained a tight upper bound on the probability of error for optimum receivers. We listed several desirable conditions which an approximating function g*(p) should satisfy in order to find a suitable upper bound. Now we present a similar set of conditions that an approximating function g**(p) should satisfy in order to find a lower bound. These conditions are

1) g**(p) should be continuous and differentiable.

2) g**(p) should be symmetrical about the point p = 0.5. This condition is desirable since g(p) is symmetrical about the point p = 0.5.

3) g**(p = 0) = g**(p = 1) = 0.

4) dg**/dp (p = 0) ≤ 1 and dg**/dp (p = 1) ≥ -1.

5) g**(p) ≤ g (p) for all values of p in the range 0 ≤ p ≤ 1.

A lower bound on the minimum probability of error can be derived in terms of the new upper bound function $g_N(p)$ discussed in Section 3.5. The first, second, and third desirable conditions for the lower bound given above are satisfied by the function $g_N(p)$. The fourth condition in conjunction with the fifth are needed to guarantee that the lower bound function is below the function g(p) for all values of p in the range 0 ≤ p ≤ 1. Now we propose a lower bound function of the form

$$g_L(p) \ = \ \varepsilon \times 0.5 \ (\sin \pi p) \ exp \ [-1.8063 \ (p-0.5)^2]$$ (3.88)

The scaling factor ε is a constant yet to be determined whose value is less than unity. With the value of ε < 1, the fourth condition is satisfied. The fifth condition is satisfied when

$$\varepsilon 0.5 \ (\sin \pi p) \ exp \ [-1.8063 \ (p-0.5)^2] \ \leq p, \ 0 \leq p \leq 0.5$$ (3.89)

Our objective is to determine those values of ε for which (3.89) is satisfied. A detailed examination of (3.89) shows that it is satisfied for the following values of ε

$$0 \le \varepsilon \le 0.79 \tag{3.90}$$

With the value of $\varepsilon = 0.79$, (3.89) is achieved with equality for a specific value of p. The point p at which this happens is p = 0.275. The tightest lower bound is achieved by using $\varepsilon = 0.79$. Using this value of ε, we see that the lower bound on the minimum probability of error is given by

$$P(E) \ge \int_x (0.395\,(\sin \pi p)\,exp\,[-1.8063\,(p-0.5)^2])\,(\pi_0 f_0(x) + \pi_1 f_1(x))\,dx \tag{3.91}$$

Comparing (3.91) and (3.71), we see that the relationship between the upper bound and the lower bound can be stated as

$$\text{Lower bound} = 0.79 \times \text{Upper bound} \tag{3.92}$$

In Fig. 3.6 we plot the functions $g_L(p)$, $g_N(p)$, and $g(p)$ as a function of p. As seen from this figure, the lower bound function closely approximates the function $g(p)$ for most of the values of p except around the point p = 0.5. The tightness of the bounds can also be expressed as the ratio of the upper bound to the lower bound. This ratio should be as close to unity as possible. For the new bounds derived in this chapter this ratio is equal to 1.265 indicating that both the upper and the lower bounds are tight bounds.

The same argument as above applies to the case of symmetrical cost assignment, i.e., a lower bound function on the cost can be defined in terms of the upper bound function $g_R(\beta_1)$ as

$$g_{RL}(\beta_1) = 0.79 g_R(\beta_1) \tag{3.93}$$
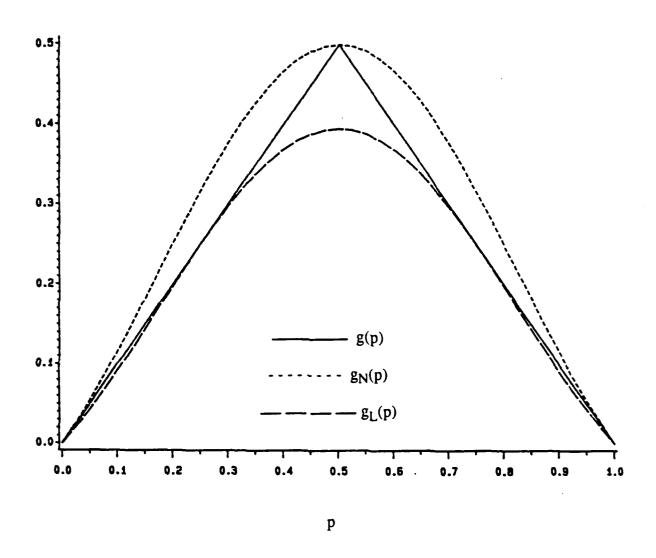
Fig. 3.6. The exact mimimum probability of error, the new upper bound, and the new lower bound functions plotted as a function of p.

The MAC can therefore be lower bounded by

$$R \geq E_x \{ g_{RL}(\beta_1(x)) \} = 0.79 \int_x g_R(\beta_1(x)) (\pi_0 f_0(x) + \pi_1 f_1(x)) dx$$

$$(3.94)$$

3.7 Example

In this section we present a numerical example where we compare the exact minimum probability of error in a hypothesis testing problem with the new upper and lower bounds on the minimum probability of error derived in this chapter. For the sake of comparison we also consider the Bhattacharyya upper and lower bounds.

Consider the hypothesis testing problem in which the observation X under H_0 is a Gaussian random variable with mean m_0 and variance σ_0^2 and when H_1 is true the observation X is a Gaussian random variable with mean m_1 and variance σ_1^2. The hypothesis testing problem can be represented as

$$H_0 : X_i \sim N(m_0, \sigma_0^2)$$

$$H_1 : X_i \sim N(m_1, \sigma_1^2)$$

We assume that $\pi_0 = \pi_1 = 0.5$. The probability of error is minimized using the likelihood ratio test given by (2.2)

$$\frac{f_1(x)}{f_0(x)} \underset{H_0}{\overset{H_1}{\gtrless}} 1 \qquad (3.95)$$

where

$$f_i(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left[-\frac{(x-m_i)^2}{2\sigma_i^2} \right], i = 0, 1 \qquad (3.96)$$

For the Gaussian density functions given by (3.96), the Bhattacharyya coefficient is given by [31]

55

$$\rho_B = exp\left\{-\left[\frac{(m_1-m_0)^2}{4(\sigma_1^2+\sigma_0^2)} + \frac{1}{2}ln\left(\frac{\sigma_1^2+\sigma_0^2}{2\sigma_1\sigma_0}\right)\right]\right\} \tag{3.97}$$

We consider two cases. In the first case we assume that $\sigma_0{}^2 = \sigma_1{}^2 = 1$, $m_0 = 0$ and m_1 is a variable. In Fig. 3.7 we plot the exact minimum probability of error as a function of m_1 when decision rule (3.95) is employed alongwith the new upper and lower bounds on the minimum probability of error determined from (3.71) and (3.91). We also plot the Bhattacharyya upper and lower bounds determined from (3.49) and (3.50). In the second case, we assume that $m_0 = 0$, $m_1 = 2$, $\sigma_0{}^2 = 1$ and $\sigma_1{}^2$ is a variable. The exact minimum probability of error, the new upper and lower bounds, and the Bhattacharyya upper and lower bound as a function of $\sigma_1{}^2$ are plotted in Fig. 3.8. Figures 3.7 and 3.8 exhibits the tightness of the new upper and lower bounds.

3.8 Summary

In this chapter we derived upper and lower bounds on the minimum probability of error and the minimum average cost for optimum receivers. Some of the bounds on the MAC are generalizations of the known bounds on the minimum probability of error in terms of the Bhattacharyya bound and the Chernoff bound. Furthermore, we derived a new upper bound on the probability of error which is tighter than the previously available bounds. This bound was generalized to the Bayesian problem with symmetrical costs. Tight Lower bounds on the probability of error and the MAC were also derived in terms of the new upper bound. These bounds are used for the performance characterization of optimum receivers. In addition, these bounds can be employed for the design of quantizers in signal detection systems, for the design of decentralized detection systems, for signal design and other related problems in communication systems.
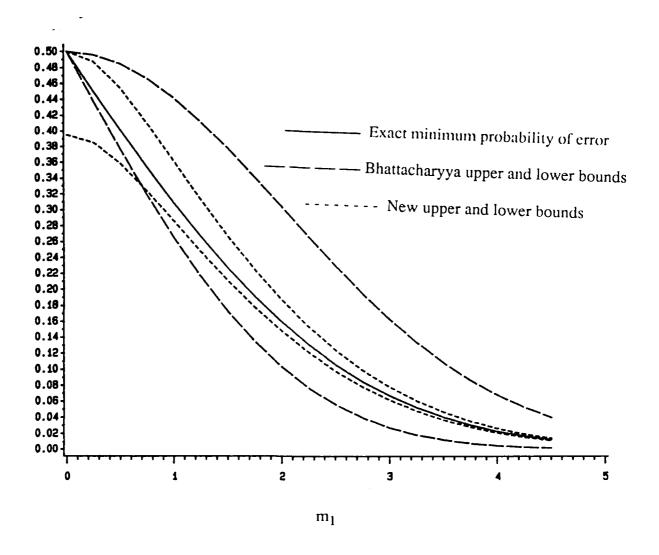
Fig. 3.7. Comparison of the new bounds, the Bhattacharyya bounds, and the exact minimum probability of error when m_1 is varied.
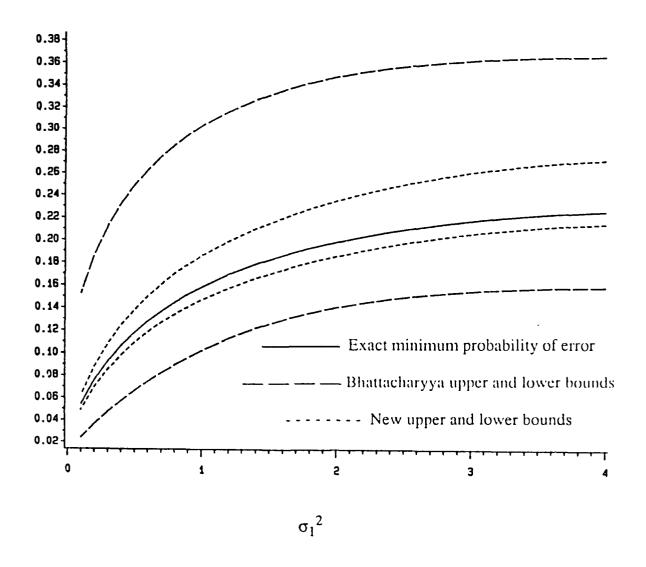
Fig. 3.8. Comparison of the new bounds, the Bhattacharyya bounds, and the exact minimum probability of error when σ_1^2 is varied.

CHAPTER 4

PERFORMANCE ANALYSIS AND DESIGN OF THE BAYESIAN PARALLEL FUSION SYSTEM

4.1 Introduction

As discussed in Chapter 1, in a decentralized detection system with data fusion, a group of local detectors process the observations they receive regarding the status of a certain phenomenon, and transmit their decisions to a fusion center where the final decision is made. The design of this system involves specifying both the local decision rules and the fusion rule. Several authors have dealt with the design of such a system using different criteria. Tenney and Sandell [7] treated the decentralized detection problem with no data fusion from a Bayesian point of view. Costs were assigned to reflect the various courses of actions of each local detector. The local decision rules were chosen such that the average cost is minimized. Assuming known sensor thresholds, Chair and Varshney [9] developed a minimum average cost algorithm for combining the sensor decisions at the fusion center. Optimization of the entire system was considered by Hoballah and Varshney [10] where they obtained a person-by-person optimal solution to the problem. The optimization of the entire system was also considered by Reibman and Nolte [12] where an exhaustive search has to be done over all the possible fusion rules in order to determine the overall minimum cost solution. Chair and Varshney [14] considered the problem of distributed Bayesian hypothesis testing with distributed data fusion in which data fusion is performed at each site.

Distance measures from the class of Ali-Silvey distance measures have recently been used for the design of decentralized detection systems [35,38]. In fact, this class has received an increasing interest in the design of quantizers for hypothesis testing because of its strong link to the probability of error (POE). Let $\pi = \{\pi_0, \pi_1\}$ be the set of all permissible pairs of the prior probabilities in a binary hypothesis testing problem. Then, there exists a subset of π for which if the distance between a given set of conditional densities is larger

than the distance between another set of conditional densities, then the POE corresponding to the first set is less than the POE corresponding to the second set. This result is known as the Blackwell theorem [31]. Poor and Thomas [32] applied the general class of Ali-Silvey distance measures to the design of a generalized quantizer for binary decision systems. Poor [33] also used these measures as criteria for analyzing the effects of fine data quantization on inferential procedures and for designing quantizers to minimize these deterimental effects. A number of authors have applied members of the class of Ali-Silvey distance measures into the design of decentralized detection systems. For example, Longo, Lookabaugh, and Gray [35] have employed the Bhattacharyya distance, whereas Lee and Chao [36] have used the J-divergence to subpartition the decision space when a quality bit is to be transmitted along with the decision to the fusion center. Kazakos [57] has employed these distance measures to obtain bounds on the performance of distributed detection systems.

In this chapter we employ the minimum average cost (MAC) as the system performance measure for the parallel fusion system shown in Fig. 1.1. Here our objectives are threefold. First, we shall examine the improvement in the system performance as a function of the number of sensors. To achieve this objective, we need an explicit relationship between the minimum average cost and the number of sensors. Second, we shall obtain optimum local thresholds which minimize the global average cost for a given number of sensors. Third, as an example, we consider the design of the minimum probability of error (MPOE) detection system where we compare the performance of the optimum system t several suboptimum systems.

The chapter is organized in six sections. Section 4.2 contains the system description, terminology, and notation. In Section 4.3, we examine a two-sensor system and derive a condition under which the performance of this system is identical to that of a single sensor. We also derive relationships for the minimum cost of an n-sensor system in terms of the sensor decisions for the hard and the soft decision cases. In Section 4.4, we make use of the results derived in Section 4.3 to design a system with minimum average cost. In Section 4.5, we consider the design of the global minimum probability of error systems. In ad-

dition, we discuss a suboptimum system design in which all the system components are MAP receivers. Then, we discuss two suboptimum systems designed based on the discrimination and the Bhattacharyya distance measures. Next, we employ the new upper bound on the minimum probability of error derived in Chapter 3 to design a nearly optimum decentralized detection system. Finally, we present numerical examples in which we compare the performance of the various suboptimum systems to that of the optimum system. We also compare the performance of the nearly optimum system to that of the optimum system. Section 4.6 contains a summary of the work reported in this chapter.

4.2 Preliminaries

Let us consider the system shown in Figure 1.1 consisting of n local detectors and a fusion center. Here, we consider the simple binary hypothesis testing problem. The hypothesis H_0 with a priori probability π_0 is tested against the alternative hypothesis H_1 with a priori probability π_1. The system receives n observations $X_1, X_2,..., X_n$ where X_i denotes the observation received by local detector LD_i. We limit our treatment to the case when X_i is a scalar observation. The generalization to the vector case is straightforward. We assume that these observations are independent and identically distributed random variables with conditional probability density functions $p_j(x_i)$; $j = 0, 1, i = 1, ..., n$. The local detectors $LD_1, ..., LD_n$ are linked to the global decision maker (also known as the fusion center) through bandlimited channels. Due to this constraint on the channel capacities, the local detectors transmit a compressed version of the observed data to the fusion center. The compressed data can be in the form of a hard decision or a soft decision. In the first case, the local decision z_i takes on one of two possible values depending upon the local decision of local detector LD_i. In the second case, the observation space of X_i is partitioned into M nonoverlapping regions. The transmitted local decision z_i, correspondingly, takes on one of M possible values. In both cases the fusion center is responsible for making the final decision.

For $M > 2$, let $T_{i0} < T_{i1} <... < T_{iM}$ be the local thresholds of local detector LD_i with $T_{i0} = -\infty$ and $T_{iM} = \infty$. Also, let $\{a_{i1}, ..., a_{iM}\}$ be the set of values that z_i may assume.

These outcomes occur with the following probabilities

$$P_{ikj} = P(z_i = a_{ik} | H_j) = \int_{T_{i(k-1)}}^{T_{ik}} p_j(x_i) dx_i, \ 1 \leq i \leq n, 1 \leq k \leq M, j = 0, 1.$$

$$(4.1)$$

We assume that the decision rules at the local detectors are identical. This means that $T_{ik} = T_k$ and $P_{ikj} = P_{kj}$ for $1 \leq i \leq n$. It should be pointed out that the theory presented here can be easily generalized to the case where local decision rules are not identical. For simplicity of presentation, we limit our discussion in this chapter to the identical local detector case. It is worth mentioning that it has been observed in [24] that for decentralized Bayesian detection, identical decision rule assumption often results in little or no loss of optimality. In the next chapter we deal with the situation where the local decision z_i is either 0 or 1 and where the observations are not identical. We also assume that the decision made by local detector LD_i is independent of the decisions made by the other local detectors. Based on these assumptions, the decisions $z_1, \ldots, z_n$ form a sequence of i.i.d generalized Bernoulli trials with parameters $(P_{1j}, P_{2j}, \ldots, P_{Mj})$ when H_j is true. These decisions are sent over bandlimited channels to the fusion center. The fusion center which is the global decision maker bases its decision on the decision vector $\underline{U} = [z_1, \ldots z_n]$. To optimally partition the observation space formed by the discrete random vector $\underline{U}$ into the decision regions corresponding to H_0 and H_1, the fusion center performs a likelihood ratio test. Our treatment in this chapter is based on the assumption that the fusion center is a minimum average cost (MAC) receiver. If C_{kj}; $k, j = 0, 1$, denotes the overall cost of deciding H_k when H_j is true then the global decision rule corresponding to this choice of the fusion center is

$$u_0 = \begin{cases} 1 & \text{if } \Lambda(\underline{U}) > \eta \\ 0 & \text{otherwise} \end{cases}$$

$$(4.2)$$

where

$$\Lambda(\underline{U}) = \frac{P(\underline{U}|H_1)}{P(\underline{U}|H_0)} = \frac{P(\underline{U} = z_1 \ldots z_n | H_1)}{P(\underline{U} = z_1 \ldots z_n | H_0)}$$

62

$$\eta = \frac{\pi_0 (C_{10} - C_{00})}{\pi_1 (C_{01} - C_{11})}$$

and $u_0 = k$ means that the fusion center decides in favor of H_k. Making use of the independence assumption between the decisions z_i, $\Lambda(\underline{U})$ can be expressed as

$$\Lambda(\underline{U}) \approx \frac{\prod\limits_{i=1}^{n} P(z_i | H_1)}{\prod\limits_{i=1}^{n} P(z_i | H_0)} \tag{4.3}$$

Using (2.29), the MAC of the system can be expressed as

$$R = R_0 - 0.5 \sum_{\underline{U}} \left| C_1 \prod\limits_{i=1}^{n} P(z_i | H_1) - C_0 \prod\limits_{i=1}^{n} P(z_i | H_0) \right| \tag{4.4}$$

where the summation is taken over all values assumed by the vector $\underline{U}=[z_1\ z_2,\ ...\ z_n]$ with each element z_i taking one of the M possible values. For the interesting case of $M = 2$, hard decisions are made at the local detectors based on the following decision rule

$$z_i = \begin{cases} 1 \text{ if } p_1(x_i)/p_0(x_i) \geq \tau \\ \\ 0 \text{ otherwise} \end{cases} \qquad i=1,2,...,n \tag{4.5}$$

where $z_i = 1$ means that H_1 is declared true and $z_i = 0$ means that H_0 is declared true by local detector LD_i. It should be pointed out that the decision rule (4.5) is a threshold test due to the independence assumption made earlier. This decision rule thus characterizes each local detector in terms of its probabilities of detection P_D and of false alarm P_F. The decision z_i has the following conditional density functions under the two hypotheses

$$f_0(z_i) = P_F^{z_i} (1 - P_F)^{1-z_i} \tag{4.6}$$

$$f_0(z_i) = P_D^{z_i} (1 - P_D)^{1-z_i} \tag{4.7}$$

63

and the global likelihood function (4.3) becomes

$$\Lambda(\underline{U}) = \frac{\prod_{i=1}^{n} f_1(z_i)}{\prod_{i=1}^{n} f_0(z_i)} \qquad (4.8)$$

Based on the likelihood ratio test (4.2) and using (4.3) and (4.8), we derive in Appendix A the algorithm based on which the fusion center combines the decisions received from the local detectors in order to make the global decision.

At this point we need to make a few remarks regarding (4.2) and (4.4).

1) The specific values of $\underline{U}$ that satisfy equality in the likelihood ratio test (4.2) can be assigned to either one of the decision regions corresponding to H_0 and H_1 without affecting the MAC, i.e., no randomization is necessary.

2) Of all the possible fusion rules, (4.2) specifies the fusion rule that can achieve the smallest MAC for given local decision rules. Equation (4.4) specifies the cost corresponding to this (best) fusion rule. It should be pointed out that an explicit knowledge of the fusion rule is not required to determine the MAC given in (4.4). The following example which deals with the binary decision case illustrates this point.

Example 4.1

Consider the decentralized detection system of Fig. 1.1 with three local detectors. Let us consider the hypothesis testing problem in which the observations X_i are normally distributed with unit variance, zero mean when H_0 is true and unit mean when H_1 is true, i.e., we test

$H_0 : X_i \sim N(0,1)$

versus

$H_1 : X_i \sim N(1,1) \quad , i = 1, 2, 3.$

Assume that $\pi_0 = 0.5$, $C_{11} = C_{00} = 0$ and $C_{10} = C_{01} = 1$. Let us consider the case where the local detectors are designed so as to minimize their own probability of error. It can be easily seen that the probability of detection for each local detector is $P_D = 0.6915$ (or the probability of miss $P_M = 0.3085$) and the probability of false alarm $P_F = 0.3085$. The local decisions are sent to the fusion center for decision combining and for yielding a final decision. The possible fusion rules for the system are the *AND* fusion rule, the *OR* fusion rule, and the majority logic rule. The probability of error achieved by each one of these fusion rules is given by

$$P(E)_{AND} = \pi_0 P_F^3 + \pi_1(3P_M - 3P_M^2 + P_M^3) = 0.41335$$

$$P(E)_{MAJ} = \pi_0(3P_F^2 - 2P_F^3) + \pi_1(3P_M^2 - 2P_M^3) = 0.22679$$

$$P(E)_{OR} = \pi_0(3P_F - 3P_F^2 + P_F^3) + \pi_1 P_M^3 = 0.28534$$

As will be seen later, applying equation (4.4), the minimum probability of error can be computed to be

$$P(E) = 0.5 - 0.5 \sum_{k=0}^{3} \binom{3}{k} \left| \pi_0 P_F^k (1 - P_F)^{3-k} - \pi_1 P_D (1 - P_D)^{3-k} \right|$$

which results in

$$P(E) = 0.22679 \qquad \blacksquare$$

As can be seen from Example 4.1, equation (4.4) resulted in the probability of error (POE) corresponding to the best fusion rule, namely, the majority logic. As illustrated in Section 4.4, this fact can be used to drastically reduce the complexity of the design of optimum decentralized detection systems.

4.3 Performance of Decentralized MAC Receivers

In this section, we investigate the performance of decentralized detection systems as

a function of number of detectors with minimum average cost as the performance measure. Initially, we consider the case of hard decisions where the local observations are quantized into two levels and the quantized value is transmitted to the fusion center. First, results are derived for a two-sensor system and then they are extended to three and n sensor systems respectively. The treatment is valid for any operating point (P_D, P_F) on the receiver operating characteristic of the local detectors, i.e., system is not necessarily the optimum system. The results are further generalized to the soft decision case, i.e., to the case with n sensors and M quantization levels at the local detectors. Once again, the results are valid for any set of local thresholds $T_0 < T_1 < ... < T_M$ which are used to perform quantization.

4.3.1 Performance of a Two-sensor System

Let us consider a two-sensor system. The fusion rules for this structure are limited to two, the AND fusion rule and the OR fusion rule (the trivial cases of always deciding in favor of H_0 or always deciding in favor of H_1 are not considered). Using (2.1) we compute the system average cost for each fusion rule as

$$R_{AND} = C_{00}\pi_0 (1 - P_F^2) + C_{10}\pi_0 P_F^2 + C_{01}\pi_1 (1 - P_D^2) + C_{11}\pi_1 P_D^2$$

(4.9)

$$R_{OR} = C_{00}\pi_0 (1 - P_F)^2 + C_{10}\pi_0 (2P_F - P_F^2) + C_{01}\pi_1 (1 - P_D)^2 + C_{11}\pi_1 (2P_D - P_D^2)$$

(4.10)

If we denote by R_2 the MAC of the two sensor system, then

$$R_2 = min (R_{AND}, R_{OR})$$

(4.11)

Substituting (4.9) and (4.10) in (4.11) and using (2.16) we get

$$R_2 = R_1 - |\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) - \pi_0 P_F (1 - P_F) (C_{10} - C_{00})|$$

$$(4.12)$$

where

$$R_1 = C_{00} \pi_0 (1 - P_F) + C_{10} \pi_0 P_F + C_{01} \pi_1 P_M + C_{11} \pi_1 (1 - P_M)$$

The first term on the right hand side of (4.12), R_1, represents the average cost for a single sensor system. The second term, which is a positive quantity, represents the improvement in performance when two sensors are used with the same cost assignment as before at the fusion center. An interesting situation arises when

$$\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) = \pi_0 P_F (1 - P_F) (C_{10} - C_{00}) \qquad (4.13)$$

In this case, the improvement term in (4.12) vanishes and, consequently, the performance of th two sensor system reduces to that of a single sensor.

4.3.2 Performance of a Three-sensor System

When the number of local detectors is three, the possible fusion rules for the system are the AND fusion rule, the MAJORITY logic fusion rule, and the OR fusion rule. These fusion rules yield the following average costs

$$R_{AND} = C_{00} \pi_0 (1 - P_F^3) + C_{10} \pi_0 P_F^3 + C_{01} \pi_1 (1 - (1 - P_M)^3) + C_{11} \pi_1 (1 - P_M)^3$$

$$(4.14)$$

$$R_{OR} = C_{00} \pi_0 (1 - P_F)^3 + C_{10} \pi_0 (1 - (1 - P_F)^3) + C_{01} \pi_1 P_M^3 + C_{11} \pi_1 (1 - P_M^3)$$

$$(4.15)$$

$$R_{MAJ} = C_{00}\pi_0 (1 - 3P_F^2 + 2P_F^3) + C_{10}\pi_0 (3P_F^2 - 2P_F^3) + C_{01}\pi_1 (3P_M^2 - 2P_M^3)$$

$$+ C_{11}\pi_1 (1 - 3P_M^2 + 2P_M^3) \tag{4.16}$$

As in the two sensor case, the MAC for this system is

$$R_3 = min \, (min \, (R_{AND}, R_{OR}), R_{MAJ}) \tag{4.17}$$

Substituting (4.14), (4.15), and (4.16) in (4.17) we get (see Appendix B for details)

$$R_3 = R_1 - \frac{1}{4} [B + 3|A| + 3|B - |A||] = R_1 - \Delta_{13} \tag{4.18}$$

where B and |A| are given by

$$B = \pi_0 P_F (1 - 3P_F + 2P_F^2) (C_{10} - C_{00}) + \pi_1 P_M (1 - 3P_M + 2P_M^2) (C_{01} - C_{11})$$

$$|A| = |\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) - \pi_0 P_F (1 - P_F) (C_{10} - C_{00})|$$

The term between brackets on the right hand side of (4.18) represents the improvement that the three-sensor system has over the single-sensor system. Now we show that this term is always positive indicating that an improvement in system performance is guaranteed as we go from a single-sensor system to a three-sensor system.

We start first with the interesting case where $0 < P_F < 0.5$ and $0 < P_M < 0.5$. Now consider the function

68

$$u(x) = 1 - 3x + 2x^2 \tag{4.19}$$

This function is plotted in Fig. 4.1. As is clear from this figure, the function u(x) is positive for values of x in the range $0 < x < 0.5$. By examining the expression for B in equation (4.18), we see that both the first and the second term contain functions of P_F and P_M of the form (4.19). Therefore, for values of P_F and P_M in the ranges specified above, the value of B is positive. When B is positive, all three terms between brackets on the right hand side of (4.18) are positive, indicating that the improvement term Δ_{13} is positive and, therefore, it is obvious that $R_3 < R_1$. For other values of P_F and P_M that make B negative, we make the change of variables B = - D in which D is positive. Now consider the quantity Δ_{13} given in (4.18) that represents the improvement in performance of the three-sensor system over that of the single-sensor system. In terms of |A| and D, the improvement Δ_{13} can be written as

$$\Delta_{13} = R_1 - R_3 = \frac{1}{4} [3|D + |A|| - D + 3|A|] \tag{4.20}$$

But

$$|D + |A|| = D + |A| \tag{4.21}$$

Making use of (4.21), we see that

$$\Delta_{13} = \frac{1}{2} [D + 3|A|] \tag{4.22}$$

which clearly shows that $\Delta_{13} \geq 0$. To examine whether a condition similar to (4.13) exists in which the performance of the three-sensor system is the same as the performance of a single-sensor system, we need the improvement term Δ_{13} in (4.22) to be equal to 0. This
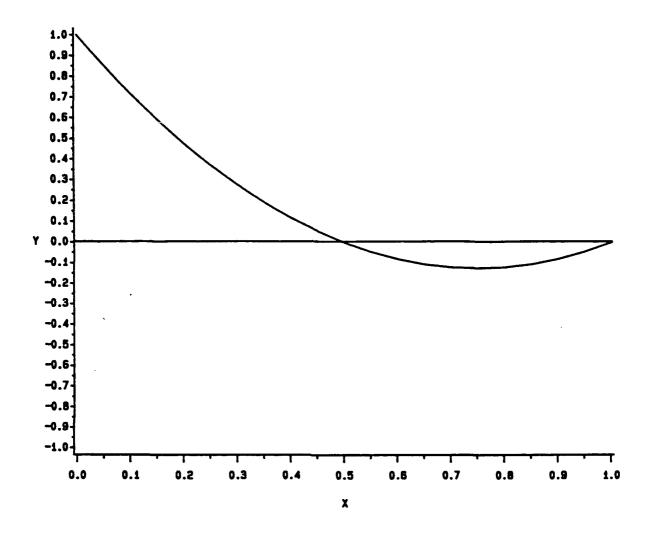
Fig. 4.1. The function $y = u(x) = 1 - 3x + 2x^2$

happens when $|A| = 0$ and $D = -B = 0$. The condition $|A| = 0$ leads to the requirement that

$$\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) = \pi_0 P_F (1 - P_F) (C_{10} - C_{00}) \qquad (4.23)$$

while the condition $B = 0$ requires that

$$\pi_0 P_F (1 - 3P_F + 2P_F^2) (C_{10} - C_{00}) = -\pi_1 P_M (1 - 3P_M + 2P_M^2) (C_{01} - C_{11})$$
$$(4.24)$$

Equation (4.24) can also be rewritten as

$$\pi_0 P_F (1 - P_F) (1 - 2P_F) (C_{10} - C_{00}) = -\pi_1 P_M (1 - P_M) (1 - 2P_M) (C_{01} - C_{11})$$
$$(4.25)$$

Substituting (4.23) into (4.25) we get

$$\pi_0 P_F (1 - P_F) (1 - 2P_F) (C_{10} - C_{00}) = -\pi_0 P_F (1 - P_F) (1 - 2P_M) (C_{10} - C_{00})$$
$$(4.26)$$

By cancelling common terms on both sides of (4.26) we obtain

$$(1 - 2P_F) = -(1 - 2P_M) \qquad (4.27)$$

Simplifying (4.27) we get a relationship between P_F and P_D as

$$P_F = 1 - P_M = P_D \qquad (4.28)$$

Therefore, a simultaneous solution to (4.23) and (4.24) exists only when $P_F = P_D$ which by

71

the virtue of (4.23) requires that

$$\pi_1 (C_{01} - C_{11}) = \pi_0 (C_{10} - C_{00}) \tag{4.29}$$

As is well known from detection theory [1], the requirement $P_F = P_D$ in (4.28) contradicts the concavity property of receiver operating characteristics of optimum receivers. Therefore, a simultaneous solution to (4.23) and (4.24) does not exist and, consequently, we conclude that the improvement term in (4.18) and (4.20) cannot vanish for all values of P_D in the range $0 < P_D < 1$ and P_F in the range $0 < P_F < 1$ and a three-sensor system is always superior to a single sensor system.

It is also of interest to compare the performance of the three-sensor system to the performance of the two-sensor system. In terms of the quantities $|A|$ and B introduced in (4.18), we can express (4.12) as

$$R_2 = R_1 - |A| \tag{4.30}$$

Denote by Δ_{23} the improvement in performance of the three-sensor system over the two-sensor system. Subtracting (4.18) from (4.30) we get

$$\Delta_{23} = R_2 - R_3 = \frac{1}{4} [3||A| - B| - (|A| - B)] \tag{4.31}$$

It is evident from (4.31) that $\Delta_{23} \geq 0$. The improvement term Δ_{23} vanishes when $|A| = B$. This happens in two cases

Case 1. A = B

Using the values of $|A|$ and B defined in (4.18) we get the condition

$$\frac{\pi_1 (C_{01} - C_{11})}{\pi_0 (C_{10} - C_{00})} = \frac{P_F (1 - P_F)^2}{P_D (1 - P_D)^2} \tag{4.32}$$

Case 2. -A = B

This condition leads to the requirement

$$\frac{\pi_1 (C_{01} - C_{11})}{\pi_0 (C_{10} - C_{00})} = \frac{P_F^2 (1 - P_F)}{P_D^2 (1 - P_D)} \tag{4.33}$$

We conclude that if either condition (4.32) or (4.33) is satisfied, then, the performance of the three-sensor system is identical to the performance of the two-sensor system. Otherwise, the three-sensor system outperforms the two-sensor system.

4.3.3 Performance of an n-sensor System

For n > 3 listing all possible fusion rules and repeated application of (2.16) to find the minimum average cost of the system using the best fusion rule becomes tedious. However, we are able to find the minimum average cost of the system with the best fusion rule in an alternative way. Recall that z_i, the decision of local detector LD_i has a Bernoulli distribution with parameter P_F when H_0 is true and parameter P_D when H_1 is true. Under our assumption of independence and identical sensors, the vector $\underline{U}$ is a sequence of Bernoulli trials. If K represents the number of sensors that decide in favor of H_1, i.e., the number of 1's in the vector $\underline{U}$, then K has the following conditional distributions

$$P_0(k) = \binom{n}{k} P_F^k (1 - P_F)^{n-k} \tag{4.34}$$

$$P_1(k) = \binom{n}{k} P_D^k (1 - P_D)^{n-k} \tag{4.35}$$

The system minimum average cost based on the vector $\underline{U}$ is found with the aid of (2.22) as

$$R_n = R_0 - 0.5 \sum_{k=0}^{n} \binom{n}{k} \left| C_1 P_D^k (1 - P_D)^{n-k} - C_0 P_F^k (1 - P_F)^{n-k} \right| \qquad (4.36)$$

Equation (4.36) is an explicit relationship between the system MAC and the number of sensors. It also represents the minimum system average cost corresponding to the best fusion rule implemented by the fusion center. Let m be an integer larger than n, then the quantity R_n - R_m represents the improvement in system performance when (m-n) new sensors are added to a system originally having n sensors.

4.3.4 Performance of an M-level Quantized System

In the preceding subsection, the performance evaluation was carried out for the case when LD_i transmits a single hard decision to the fusion center. In that case the probabilities of detection and false alarm can be appropriately defined. In this section we will generalize the results to the case where local detector LD_i transmits one of the M possible soft decisions to the fusion center.

Recall from Section 4.1 that for the case M > 2, the vector $\underline{U}$ is a sequence of n i.i.d generalized Bernoulli trials. Let X_k, $1 \le k \le M$, represent the number of sensors that have decided in favor of symbol a_k. The number of such symbols in $\underline{U}$ follows the following conditional densities

$$P(X_1 = x_1, ..., X_M = x_M | H_0) = \binom{n}{x_1 x_2 ... x_M} \prod_{k=1}^{M} P_{k0}^{x_k} \qquad (4.37)$$

$$P(X_1 = x_1, ..., X_M = x_M | H_1) = \binom{n}{x_1 x_2 ... x_M} \prod_{k=1}^{M} P_{k1}^{x_k} \qquad (4.38)$$

where

$$\binom{n}{x_1 x_2 \dots x_M} = \frac{n!}{x_1! x_2! \dots x_M!}$$

$$\sum_{k=1}^{M} x_k = n$$

and P_{kj}; $j=0,1$ in the case of identical sensors are given by (4.1). The MAC for this system is found using (2.21) as

$$R_M(n) = R_0 - 0.5 \sum_{x_1, \dots, x_M} \binom{n}{x_1 x_2 \dots x_M} \left| C_1 \prod_{k=1}^{M} P_{k1}^{x_k} - C_0 \prod_{k=1}^{M} P_{k0}^{x_k} \right| \qquad (4.39)$$

The summation on the right hand side of (4.39) is taken over all possible values of x_1, x_2, ..., x_M such that $x_1+x_2+...+x_M = n$. Analogous to the case $M = 2$, the quantity $R_M(n) - R_M(m)$ represents the improvement in the system performance when (m-n) new local detectors are added to a system originally composed of n sensors.

4.4 Design of the Optimum Decentralized MAC Detection System

The two major problems encountered in the design of a decentralized detection system with data fusion for a given number of local detectors are the determination of the optimum fusion rule and the optimum local decision rules. For the first part of the design problem, we assume the fusion center to be a minimum average cost receiver with (4.2) as the decision rule. This decision rule when implemented at the fusion center optimally partitions the observation space of the decision vector $\underline{U}$ into two mutually exclusive decision regions corresponding to hypotheses H_0 and H_1. This partitioning yields the smallest achievable average cost for any given arbitrary local decision rules. The partitionning corresponds to a specific fusion rule (see Appendix A for the algorithm used by the fusion center to combine the local decisions and to make the global decision). In the previous section we derived the expression for the minimum average cost for an n-sensor system

with two quantization levels as well as M quantization levels. For the second part of the design problem, our objective is to select the local decision rules such that the average cost is globally minimized resulting in the optimum decentralized MAC system.

Let $\underline{T} = [T_1, T_2, ..., T_{M-1}]$ be the vector consisting of local thresholds at any of the identical detectors. The probabilities P_{kj}, $j = 0, 1$, given by (4.1) are functions of the thresholds. We stress this by expressing P_{k0} and P_{k1} as

$$P_{kj} = P_{kj}(T_k, T_{k-1}), j = 0, 1 \qquad (4.40)$$

for $1 \leq k \leq M$. For a given threshold vector $\underline{T}$, the global decision rule (4.2) expressed as an explicit function of $\underline{T}$ becomes

$$u_0(\underline{T}) = \left\{ \begin{array}{ll} 1 & \text{if } \Lambda(\underline{T}) > \eta \\ \\ 0 & \text{otherwise} \end{array} \right. \qquad (4.41)$$

where

$$\Lambda(T) = \frac{\displaystyle\prod_{i=1}^{n} P(z_i(T)|H_1)}{\displaystyle\prod_{i=1}^{n} P(z_i(T)|H_0)}$$

For a given value of $\underline{T}$, the decision rule (4.41) yields the smallest average cost among all other decision rules (fusion rules). The resulting system MAC as a function of $\underline{T}$ can be expressed as

$$R_{Mn}(T) = R_0 - 0.5 \sum_{X_1, ..., X_M} \left(x_1 x_2 ... x_M \right) \left| C_1 \prod_{k=1}^{M} P_{k1}^{x_k}(T_k, T_{k-1}) - C_0 \prod_{k=1}^{M} P_{k0}^{x_k}(T_k, T_{k-1}) \right|$$

It should be reemphasized that, for a fixed $\underline{T}$, if one evaluates the system average cost for all possible fusion rules, then (4.42) gives the minimum among all of the above system av-

erage costs and the fusion rule which yields the minimum is specified by (4.41). For a different value of $\underline{T}$, say $\underline{T}''$, the entire process needs to be repeated, i.e., equations (4.42) and (4.41) provide the minimum system average cost for the local threshold vector T'', and the corresponding fusion rule respectively. Therefore, to design the overall system, $R_{Mn}(\underline{T})$ should be minimized with respect to $\underline{T}$ and the corresponding fusion rule can be determined from (4.41). The resulting system, i.e., the resulting value of $\underline{T}$ and the corresponding fusion rule, yields the minimum achievable cost for the system.

Note that the above design procedure requires the minimization of a single function of (M-1) variables. Also, even when the assumption of identical thresholds at the local detectors is relaxed, we need to minimize a single function of n(M-1) variables. This means that for the binary case (M=2), we need to minimize a function of n variables only. This case of non identical thresholds will be considered in the next chapter. Our procedure is computationally simpler than the design procedures available in the literature where optimization needs to be carried out for all possible fusion rules. Two methods that deal with the design of binary decentralized Bayesian detection systems are reported in the literature. In the first method [12], the fusion center is fixed and a set of n coupled nonlinear equations are solved to determine the n local thresholds. This has to be repeated for all the permissible fusion rules. The solution with the smallest overall cost is finally selected as the optimum system. The exponential growth of the number of fusion rules to be searched makes the use of this method impractical. The other method is the person-by-person optimization procedure [10] where conditions are determined for each local detector and the fusion center so as to minimize the system cost when the other system components are assumed to be fixed. The resulting equations are solved simultaneously to yield the person-by-person optimal solution. This solution is not necessarily the optimum solution. This procedure requires a simultaneous solution of (2^n+n) coupled nonlinear equations for the binary hypothesis testing problem. Thus, our design procedure is computationally efficient and yields the optimum solution. It can be employed for the design and performance evaluation of relatively large detection networks as shown in the next chapter.

In our design procedure when M = 2, the n-dimensional vector $\underline{T}$ reduces to a scalar.

In this case the local detectors perform their likelihood ratio tests with respect to a common threshold τ (see equation 4.5). For a given value of τ, the local detection probability P_D and the local false alarm probability P_F are functions of τ. We emphasize this by expressing P_D as $P_D(\tau)$ and P_F as $P_F(\tau)$. Therefore, we can express the conditional distributions given by (4.6) and (4.7) of the local decision z_i in terms of τ as

$$f_0(z_i(\tau)) = [P_F(\tau)]^{z_i} [1-P_F(\tau)]^{1-z_i} \tag{4.43}$$

$$f_1(z_i(\tau)) = [P_D(\tau)]^{z_i} [1-P_D(\tau)]^{1-z_i} \tag{4.44}$$

The conditional densities of the decision vector $\underline{U}$ correspondingly can be expressed as

$$f_j(\underline{U}(\tau)) = \prod_{i=1}^{n} f_j(z_i(\tau)), j = 0, 1 \tag{4.45}$$

The likelihood function in (4.41) expressed as a function of τ is

$$\Lambda(\underline{U}(\tau)) = \frac{f_1(\underline{U}(\tau))}{f_0(\underline{U}(\tau))} \tag{4.46}$$

As pointed out in remark 2 in Section 4.2, the decision rule (4.8) with $\Lambda(\underline{U}(\tau))$ as given in (4.46) specifies the fusion rule with the smallest MAC for a given value of τ. The MAC corresponding to this fusion rule with n incoming decisions expressed in terms of τ is

$$R_n(\tau) = R_0 - \frac{1}{2} \sum_{k=0}^{n} \binom{n}{k} \left| C_1 P_D^k(\tau) (1 - P_D(\tau))^{n-k} - C_0 P_F^k(\tau) (1 - P_F(\tau))^{n-k} \right| \tag{4.47}$$

To obtain the system with global optimum cost (least achievable MAC), $R_n(\tau)$ should be minimized with respect to the threshold τ. The resulting local threshold alongwith the corresponding fusion rule specified by (4.46) and (4.41) will yield the optimum system.

4.5 Minimum Probability of Error Systems

In this section we consider the design of the global minimum probability of error (MPOE) systems which is a special case of the results obtained in the previous section. In addition, we consider the design of several suboptimum decentralized detection systems and compare their performances. The first suboptimum system is based on local optimization in that each system component is an MAP receiver. The next two suboptimum systems are based on the discrimination and the Bhattacharyya distance measures. While these two measures have been utilized before in the design of decentralized detection systems [35,38], here we employ them in a different manner. Finally, we present a suboptimum design procedure based on our new tight upper bound on the probability of error developed in Chapter 3. For simplicity, we restrict our attention to the case $M = 2$ in this section. We conclude this section with two numerical examples that compare the performance of the optimum system to the performances of the various suboptimum systems.

4.5.1 Global MPOE System Design

As an important application of the general Bayesian formulation developed in the previous section, we consider the cost assignment $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$. This corresponds to the minimum probability of error criterion which is widely used. When the decentralized system is to be designed using the probability of error as the system performance measure, the expression for the minimum probability of error as a function of the local threshold τ is given by

$$P_E(\tau) = 0.5 - \frac{1}{2} \sum_{k=0}^{n} \binom{n}{k} \left| \pi_1 P_D^k(\tau) (1 - P_D(\tau))^{n-k} - \pi_0 P_F^k(\tau) (1 - P_F(\tau))^{n-k} \right|$$

(4.48)

To design the global MPOE system, we need to determine the value of τ that minimizes $P_E(\tau)$ given by (4.48) and then determine the fusion rule which corresponds to that value of the probability of error using (4.41) (see also Appendix A).

4.5.2 System Design Based on Local Optimization

In this section we assume that the strategy of the local detectors is to select the threshold that minimizes the POE at the local level. Let $p_0(x_i)$ and $p_1(x_i)$ be the conditional densities of the observation X_i received by local detector LD_i. The decision rule (4.5) that determines the decisions z_i becomes

$$z_i = \begin{cases} 1 \text{ if } p_1(x_i)/p_0(x_i) \geq \pi_0/\pi_1 \\ \\ 0 \text{ otherwise} \end{cases} \qquad (4.49)$$

It is clear from (4.49) that the threshold $\tau_L = \pi_0/\pi_1$ and the system POE is $P_E(\tau_L)$. When decision rule (4.49) is implemented at the local level, all the system components in Fig.1.1 become MAP receivers. Even though each component is individually optimized, the system as a whole can be far from optimum, since the sensor decisions are made independently of the global decision.

4.5.3 System Design Based on Discrimination

Here, we describe the design procedure based on the discrimination measure. The discrimination between two conditional probability density functions $f_0(x)$ and $f_1(x)$ of a continuous random variable X is defined as

$$J = \int_x f_0(x) \, ln\frac{f_0(x)}{f_1(x)} dx \qquad (4.50)$$

When the random variable X is discrete, the integration in (4.50) is replaced by a summation and the probability density functions are replaced by probability mass functions. In this case the discrimination J becomes

$$J = \sum_x P(X = x|H_0) \, ln\frac{P(X = x|H_0)}{P(X = x|H_1)} \qquad (4.51)$$

Several important properties of the discrimination measure are listed in [67]. Since the fi-

nal decision in the decentralized detection system is based on the vector $\underline{U}=[z_1,z_2,...,z_n]$, we will examine the discrimination between the conditional distributions of the vector $\underline{U}$ under the two hypotheses. This is given by

$$J_n(P_D, P_F) = \sum_{\underline{U}} P(\underline{U} = z_1...z_n | H_0) \, ln \frac{P(\underline{U} = z_1...z_n | H_0)}{P(\underline{U} = z_1...z_n | H_1)} \qquad (4.52)$$

The local decisions z_i are Bernoulli random variables with conditional density functions given by (4.43) and (4.44). Recognizing that $\underline{U}$ is a sequence of n independent and identical observations, then using the additivity property of the discrimination measure we get

$$J_n(P_D(\tau), P_F(\tau)) = nJ_1(\tau) \qquad (4.53)$$

Here, $J_1(\tau)$ is the discrimination between the conditional distributions of a single local decision z_i under the two hypotheses. Using (4.43) and (4.44) we get

$$J_1(\tau) = P_F(\tau) \, ln \frac{P_F(\tau)}{P_D(\tau)} + (1 - P_F(\tau)) \, ln \frac{(1 - P_F(\tau))}{(1 - P_D(\tau))} \qquad (4.54)$$

As an explicit function of the local threshold τ, the discrimination between the conditional densities of $\underline{U}$ can be expressed as

$$J_n(\tau) = n \left[P_F(\tau) \, ln \frac{P_F(\tau)}{P_D(\tau)} + (1 - P_F(\tau)) \, ln \frac{(1 - P_F(\tau))}{(1 - P_D(\tau))} \right] \qquad (4.55)$$

By making use of the Blackwell theorem we observe that an indirect way of minimizing the POE is to maximize the discrimination. Denote by τ_D the specific value of the local threshold τ that maximizes $J_n(\tau)$. The system POE in this case is $P_E(\tau_D)$ where $P_E(\tau)$ is given by (4.48). Two factors make this system suboptimum. First, The discrimination as defined in (4.50) is independent of the prior probabilities. This fixes the local thresholds

81

and consequently the values of P_F and P_D to be used during the system design. Designing the system based on this measure when the prior probabilities are known, in fact, means that part of the information available about the system is not exploited. Second, the discrimination measure is independent of the fusion rule and, therefore, it is not a part of the overall optimization process. This leads to the suboptimum system.

4.5.4 System Design Based on Bhattacharyya Distance

Now, we consider the design procedure based on the Bhattacharyya distance. Let X_1 be the random observation in a hypothesis testing problem with corresponding conditional densities $f_0(x_1)$ and $f_1(x_1)$. The Bhattacharyya coefficient is given by

$$\rho_1 = \int_{x_1} \sqrt{f_0(x_1) f_1(x_1)}\, dx_1 \tag{4.56}$$

When the random variable X_1 is discrete, the Bhattacharyya coefficient is defined as

$$\rho_1 = \sum_{x_1} \sqrt{P(X_1 = x_1 | H_0) P(X_1 = x_1 | H_1)} \tag{4.57}$$

The Bhattacharyya distance (see Chapter1) is defined as

$$D_1 = -\ln \rho_1 = -\ln \int_{x_1} \sqrt{f_0(x_1) f_1(x_1)}\, dx_1 \tag{4.58}$$

If $\underline{X}$ is a sequence of n independent and identically distributed observations, then the Bhattacharyya coefficient and the Bhattacharyya distance corresponding to the conditional densities of $\underline{X}$ become

$$\rho_n = \rho_1^n \tag{4.59}$$

$$\rho_n = -nln \; \rho_1 \qquad (4.60)$$

The Bhattacharyya coefficient for the conditional densities of $\underline{U}$ is

$$\rho_n (P_F, P_D) = \sum_{\underline{U}} \sqrt{(P(\underline{U} = z_1 ... z_n | H_0))(P(\underline{U} = z_1 ... z_n | H_1))} \qquad (4.61)$$

The Bhattacharyya distance between the conditional densities of the local decision z_i is

$$D_1(P_F, P_D) = -ln\,[\sqrt{P_F P_D} + \sqrt{(1 - P_F)(1 - P_D)}\,] \qquad (4.62)$$

Expressing P_D as $P_D(\tau)$ and P_F as $P_F(\tau)$ and making use of (4.60) we can express the Bhattacharyya distance between the conditional densities of $\underline{U}$ in terms of τ as

$$D_n(\tau) = -nln\,[\sqrt{P_F(\tau)P_D(\tau)} + \sqrt{(1 - P_F(\tau))(1 - P_D(\tau))}\,] \qquad (4.63)$$

Again, making use of the Blackwell theorem we see that an indirect way of minimizing the POE is to maximize the Bhattacharyya distance. Denote by τ_B the specific value of τ that maximizes $D_n(\tau)$. The system POE in this case is $P_E(\tau_B)$. Again, the resulting system is suboptimum and the reasons for this are the same ones as discussed in the previous subsection.

4.5.5 System Design Based on the New Upper Bound on the MPOE

In this section we apply the new upper bound derived in Chapter 3 to the design of decentralized detection systems. Let K be the random variable representing the number of local detectors that decide in favor of hypothesis H_1. Then K has the conditional density functions given by (4.34) and (4.35) under hypotheses H_0 and H_1 respectively. Since K is a discrete random variable then, in order to find an upper bound on the probability that the

fusion center makes an error, we need to express the new upper bound given by (3.71) in a discrete form. This is given as

$$P(E) \leq \sum_{k=0}^{n} (0.5 \, (\sin \pi p_k) \, exp \, [-1.8063 \, (p_k - 0.5)^2]) \, (\pi_0 P_0(k) + \pi_1 P_1(k))$$

$$(4.64)$$

where

$$p_k = \frac{\pi_0 P_0(k)}{\pi_0 P_0(k) + \pi_1 P_1(k)}, k = 0, 1, ..., n$$

is the a posteriori probability of hypothesis H_0 given that $K = k$. Substituting (4.34) and (4.35) into the above expression for p_k, we get

$$p_k = \frac{\pi_0 P_F^k (1 - P_F)^{n-k}}{\pi_0 P_F^k (1 - P_F)^{n-k} + \pi_1 P_D^k (1 - P_D)^{n-k}}, k = 0, 1, ..., n \qquad (4.65)$$

The upper bound in (4.64) applies for any arbitrary point (P_F, P_D) on the receiver operating characteristic of the local detectors, i.e., for any value of the local threshold. Let τ be the common threshold of the local detectors (see equation (4.5)). Expressing P_F as $P_F(\tau)$ and P_D as $P_D(\tau)$, we can write the a posteriori probabilities p_k in terms of τ as

$$p_k(\tau) = \frac{\pi_0 (P_F(\tau))^k (1 - P_F(\tau))^{n-k}}{\pi_0 (P_F(\tau))^k (1 - P_F(\tau))^{n-k} + \pi_1 (P_D(\tau))^k (1 - P_D(\tau))^{n-k}}$$

$$(4.66)$$

The upper bound on the MPOE in terms of τ becomes

$$P(E) \leq \sum_{k=0}^{n} 0.5 \binom{n}{k} \sin \pi p_k(\tau) \, (exp \, [-1.8063 \, (p_k(\tau) - 0.5)^2])$$

$$\text{x} \qquad [\pi_0 (P_F(\tau))^k (1 - P_F(\tau))^{n-k} + \pi_1 (P_D(\tau))^k (1 - P_D(\tau))^{n-k}]$$

$$(4.67)$$

The design of the system using the new upper bound calls for determining the local threshold τ^* which minimizes the right hand side of (4.67). Therefore, a suboptimum system that is designed to minimize the upper bound on the probability of error is one in which the local detectors use the threshold τ^*. The resulting POE is given by $P_E(\tau^*)$. Because of the tightness of the new upper bound, systems using this bound as the design criterion are nearly optimum. Example 4.3 demonstrates the utility of the new upper bound in the design of decentralized detection systems. Note that the knowledge of the prior probabilities is used in this approach but it is still independent of the fusion rule.

4.5.6 Examples

In this subsection, we present two numerical examples where the system performances achieved by using the different design approaches is compared.

Example 4.2

Consider a two sensor decentralized detection system used for the detection of a constant signal of level m embedded in a zero-mean Gaussian noise with variance σ^2. The null hypothesis corresponding to noise alone and the alternative hypothesis corresponding to signal plus noise are expressed as

$$H_0 : X_i \sim N(0, \sigma^2)$$

$$H_1 : X_i \sim N(m, \sigma^2)$$

For this binary hypothesis testing problem we will compare the performance of the various suboptimum systems considered previously to the performance of the optimum decentralized system as well as the optimum centralized system. Also, we will see how the performance of the optimum decentralized system improves as the number of quantization levels increases. For the one dimensional Gaussian problem the likelihood ratio test (4.5) at the local level becomes

$$\begin{array}{c} H_1 \\ p_1(x_i) \quad > \\ \overline{\qquad\qquad} \quad \tau \\ p_0(x_i) \quad < \\ H_0 \end{array} \qquad (4.68)$$

where

$$p_1(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x_i - m)^2}{2\sigma^2}\right] \qquad (4.69)$$

$$p_0(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{x_i^2}{2\sigma^2}\right] \qquad (4.70)$$

Substituting (4.69) and (4.70) into (4.68) and simplifying we get the equivalent test

$$\begin{array}{c} H_1 \\ > \\ x_i \qquad \sigma^2/m \log \tau + m/2 \; = T \\ < \\ H_0 \end{array} \qquad (4.71)$$

In this and future examples we find it more useful to evaluate the equivalent threshold T defined in (4.71) than to evaluate the threshold τ. This is because (4.71) gives us the decision regions corresponding to H_0 and H_1 in terms of the observation x, while (4.68) gives the decision regions in terms of the likelihood function. The probabilities of false alarm and detection expressed in terms of τ are given by

$$P_F(\tau) = erfc\left(\sigma\frac{\log\tau}{m} + \frac{m}{2\sigma}\right) \qquad (4.72)$$

$$P_D(\tau) = erfc\left(\sigma\frac{\log\tau}{m} - \frac{m}{2\sigma}\right) \qquad (4.73)$$

86

where erfc (u) is the complementary error function defined as

$$erfc\,(u) \;=\; \int_u^\infty \frac{1}{\sqrt{2\pi}} exp\,(-\frac{u^2}{2})\,du \qquad (4.74)$$

For parameter values $m = 1$ and $\sigma^2 = 1$, the design is carried out for all values of the a priori probability π_0. The optimum decentralized MPOE system with $M > 2$ is designed according to (4.42), while for the case $M = 2$, the system is designed according to (4.48). The suboptimum systems based on the optimization of the POE at the local levels, maximum discrimination, and maximum Bhattacharyya distance are designed according to (4.49), (4.55),and (4.63) respectively. In Fig. 4.2, we present the performance of the various suboptimum systems as compared to the optimum decentralized system and the optimum centralized system. Several observations can be drawn from this graph

1) The graph shows clearly that the optimum decentralized system significantly outperforms all three suboptimum systems over a wide range of π_0.

2) Of special interest is the point with $\pi_0 = 0.5$. At this point the local decision rule (4.49) becomes

$$z_i = \begin{cases} 1 \text{ if } p_1(x_i)/p_0(x_i) \ge 1 \\ \\ 0 \text{ otherwise} \end{cases} \qquad (4.75)$$

Using (4.75) we find that the threshold $\tau_L = 1$ and the corresponding local threshold in terms of the observations X_i is $T_L = 0.5$ (see equation (4.71)). The values of P_D and P_F are 0.6915 and 0.3085 respectively. The POE for any one of the local detectors is $P_1(E) = 0.3085$. Using (4.48) we find the POE of the two-sensor system to be $P_E(\tau = 1)$ or $P_E(T_L = 0.5) = 0.3085$. These results indicate that a single sensor system and a decentralized two-sensor system are identical in terms of performance. It is to be noted that in this case the point (P_F, P_D) satisfies condition (4.13).

3) In Fig. 4.2, we notice that over some range of π_0 the performance of the system designed based on the Bhattacharyya distance is better than that designed based on the dis

87
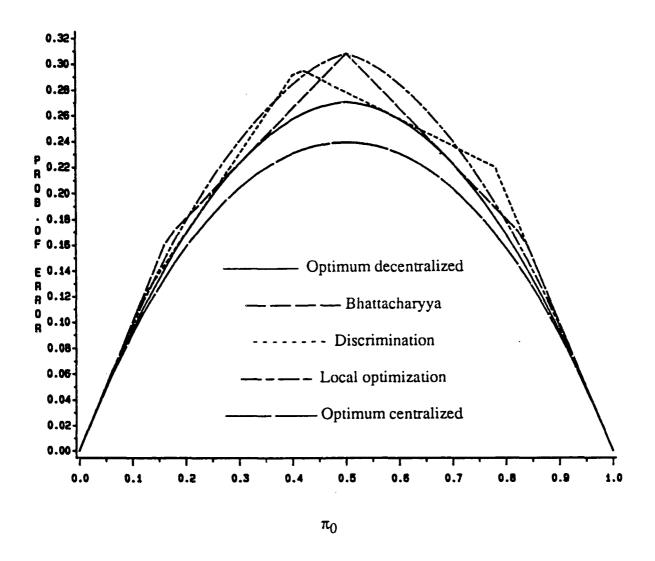
Fig. 4.2. Performance comparison of the optimum decentralized system and the suboptimum systems when each local detector processes one observation.

crimination, and over another range the opposite is true. In fact, this is consistent with the observations made in [31,32] regarding the selection of the "best" distance measure.

4) The performance curves for the systems employing maximum discrimination and maximum Bhattacharyya distance have straight line segments with different slopes. This can best be explained by substituting $C_{00} = C_{11} = 0$ and $C_{10} = C_{01} = 1$ into (4.9) and (4.10) and rewriting them as functions of π_0. The POE corresponding to the AND fusion rule and the OR fusion rule are, therefore, expressed as

$$P(E)_{AND} = (P_F^2 + P_M^2 - 2P_M) \pi_0 + (2P_M - P_M^2)$$ (4.76)

$$P(E)_{OR} = (2P_F - P_M^2 - P_F^2) \pi_0 + P_M^2$$ (4.77)

As pointed out earlier, when we design the system based on members of the class of Ali-Silvey distance measures, we find the local thresholds that maximize the distance between the conditional densities of the decision vector $\underline{U}$. These local thresholds are independent of the prior probabilities and, consequently, the values of P_F and P_D of the individual sensors are independent of these probabilities. This means that both (4.76) and (4.77) reduce to linear functions of π_0. For a given value of π_0, the fusion center computes the probability of error corresponding to each one of the fusion rules and selects the fusion rule having the smallest POE. Since the fusion center is a MPOE receiver, we see that over some range of π_0, the AND fusion rule is implemented and over another range, the OR fusion rule is implemented depending on which fusion rule has the smaller POE.

Fig. 4.3 shows the same kind of comparison considered above but assuming that each local detector processes two observations to come up with the local decision z_i. As can be seen from this figure, the relative gap between the optimum system and the suboptimum systems widens for $0 < \pi_0 < 1$. For the rest of the figures, single observations at the sensors are assumed. In Fig. 4.4, we plot the optimum local threshold given by (4.71) ver
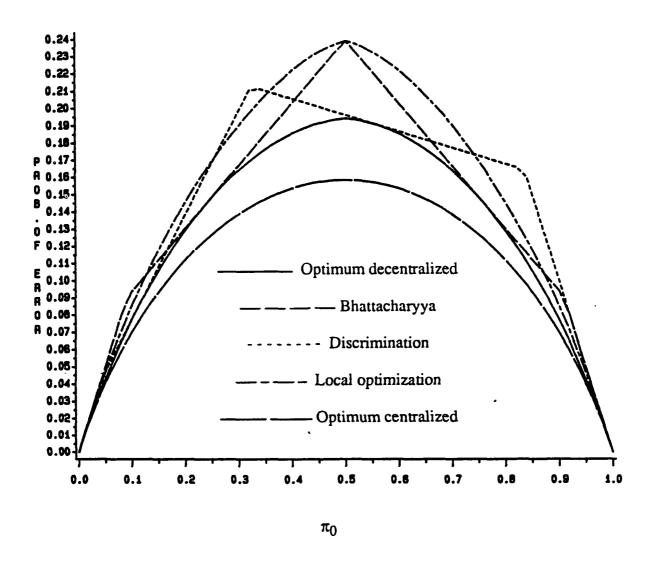
Fig. 4.3. Performance comparison of the optimum decentralized system and the subopti-
mum systems when each local detector processes two observation.
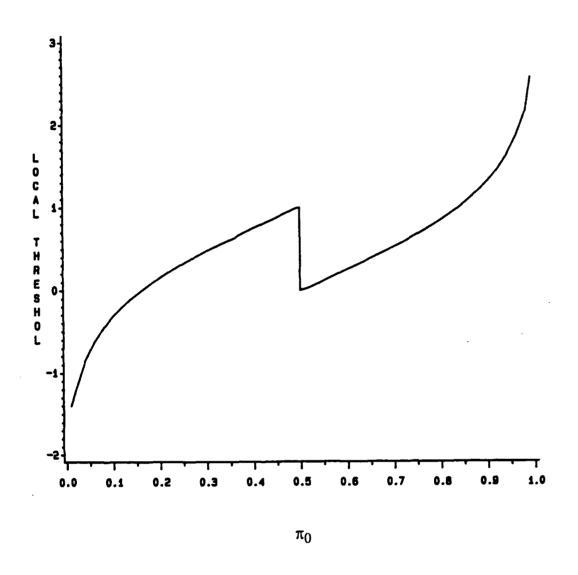
Fig. 4.4. The optimum local threshold in a two-sensor decentralized detection system.

sus π_0. In Fig. 4.5, we plot the optimum fusion rule for the optimum decentralized detection system as a function of π_0. As can be observed from Fig. 4.5, for $\pi_0 < 0.5$ the OR fusion rule performs better than the AND fusion rule and for $\pi_0 > 0.5$ the reverse is true. The abrupt change in the value of the optimum local threshold in Fig. 4.4 is a manifestation of the fact that the system always selects the fusion rule leading to the smallest cost. The variation of the optimum local threshold given by (4.71) and the optimum fusion rule as a function of π_0 for the optimum three-sensor decentralized system is shown in Figures 4.6 and 4.7 respectively. In this case the majority logic has the best performance over most of the region of π_0, i.e., for $0.095 \leq \pi_0 \leq 0.905$. In Fig. 4.8, we compare the performance of the optimum decentralized MPOE system with M = 2, 3 and 4 to the performance of the optimum centralized system. At the point $\pi_0 = 0.5$ we observe that the POE of the optimum decentralized system with M = 4 is only 2.36% higher than the POE of the optimum centralized system. In Fig. 4.9, we show the resulting MPOE as a function of the signal level m (small values of m) for the four systems when $\pi_0 = 0.5$ is assumed.

Example 4.3

The objective of this example is to design the decentralized system based on the new upper bound for the same problem as considered in Example 4.2 and compare the performance with that of the optimum decentralized detection system. We further assume that m = 1.5. The probabilities of false alarm and detection expressed in terms of τ are given by (4.72) and (4.73). The design equation for this system will be (4.67) with n = 2. In Fig. 4.10, we plot the probability of error for the system resulting from the minimization of the new upper bound as a function of π_0 (dotted curve). Also plotted in this figure is the optimum probability of error resulting from the design of the system based on equation (4.48) (solid curve). As observed from the figure, the difference between the two curves can be hardly noticed. Therefore, we also provide numerical results in Table 4.1. Due to the symmetry of the curve, numerical results for for only the values of $\pi_0 \geq 0.5$ are presented. Note that the thresholds in Table 4.1, are those specified by (4.71).
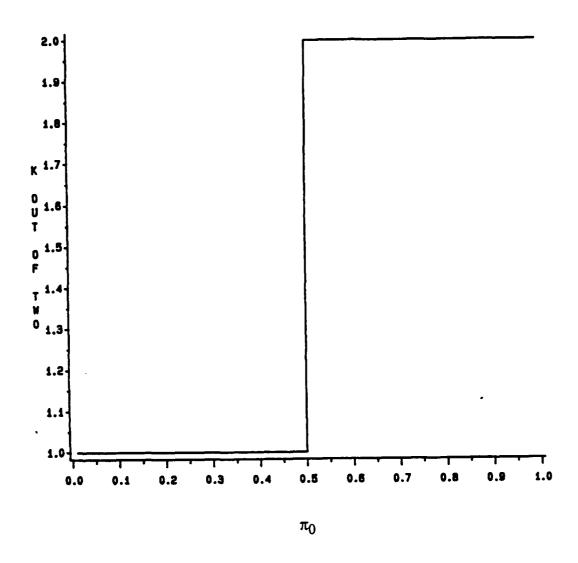
Fig. 4.5. The optimum fusion rule in a two-sensor decentralized detection system.

93
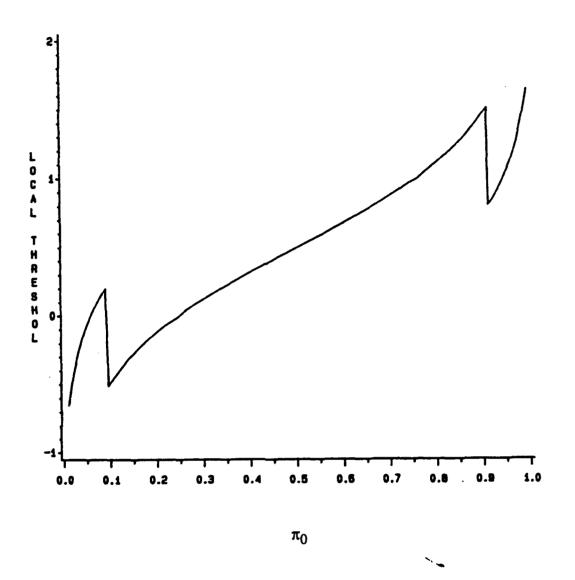
Fig. 4.6. The optimum local threshold in a three-sensor decentralized detection system.
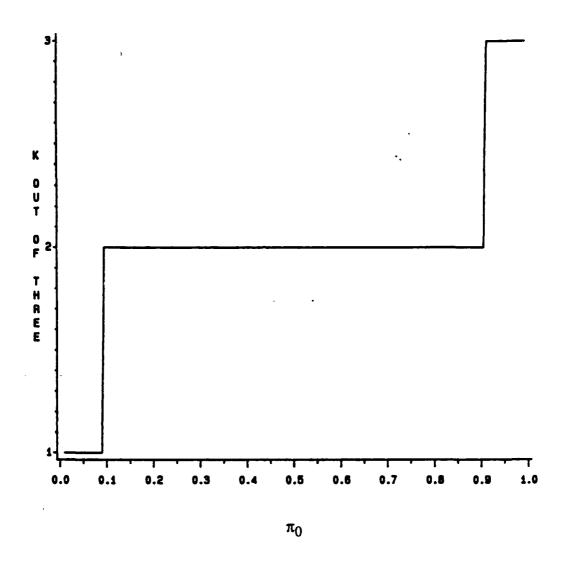
Fig. 4.7. The optimum fusion rule in a three-sensor decentralized detection system.
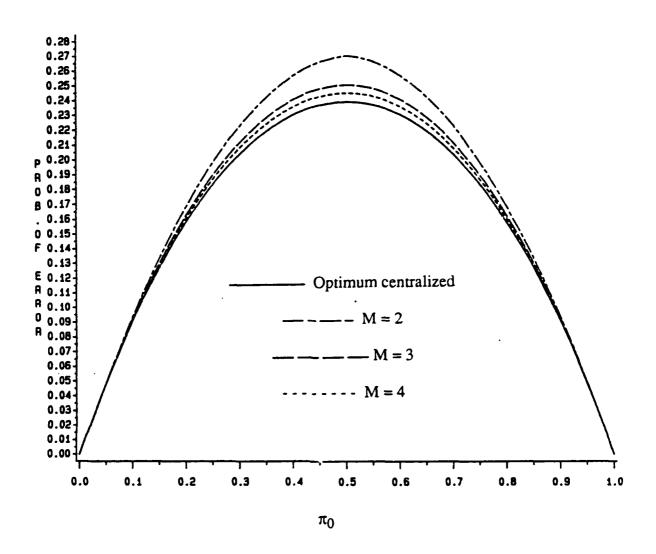
Fig. 4.8. Minimum probability of error for an M-level quantization two-sensor decentralized detection system.
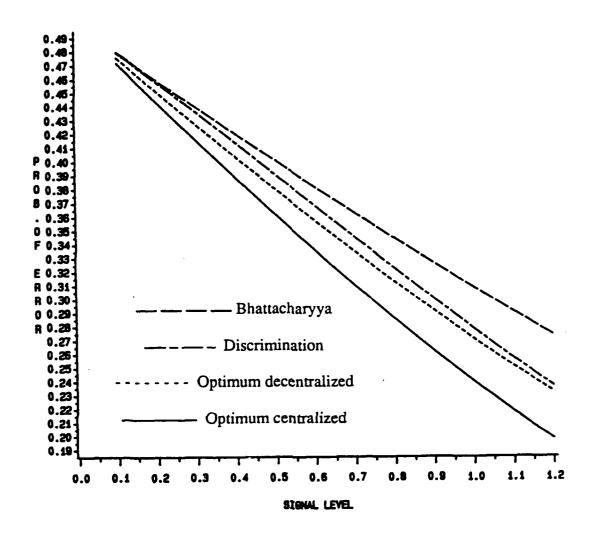
Fig. 4.9. Variation of Minimum probability of error for different systems in terms of the

signal level m when $\pi_0 = 0.5$.
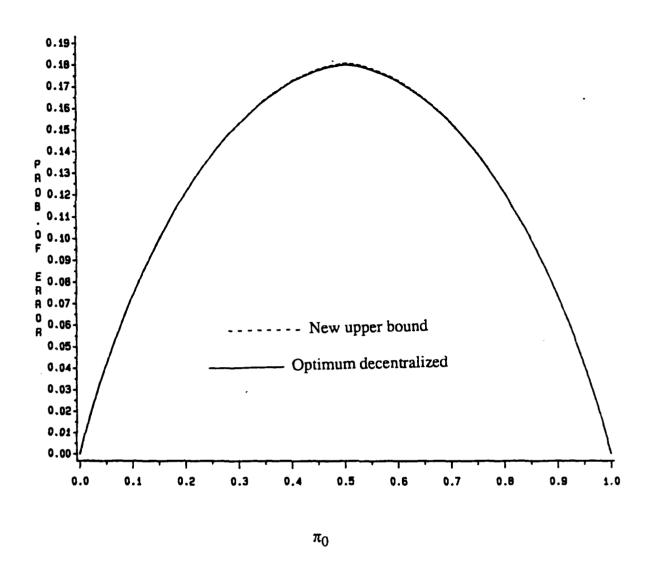
Fig. 4.10. Performance of the two-sensor system when designed based on the new upper bound as compared to that of the optimum decentralised system.

π_0	Threshold that minimizes the upper bound	POE obtained from the upper bound minimization approach	Optimum Threshold for the decentralized detection system	Optimum MPOE
0.5	0.2843	0.1810	0.2268	0.1804
0.6	0.4461	0.1726	0.4047	0.1723
0.7	0.6167	0.1529	0.5925	0.1528
0.8	0.8185	0.1209	0.815	0.1209
0.9	1.1177	0.07353	1.1375	0.0735

Table 4.1

4.6 Summary

In this chapter the design and perfc mance of minimum average cost decentralized detection system was considered.The analysis was based on the expression derived in Chapter 2 for the minimum average cost of an optimum receiver in terms of the Kolmogorov variational distance. Both hard decision and soft decision systems were considered. Performance enhancement when additional detectors are added was determined.A design approach for the MAC decentralized detection system based on the above results was presented.This computational procedure is much simpler than the previously available methods.As an example,design of distributed detection systems for MPOE criterion was considered in detail.Its performance was compared to the suboptimum systems designed based on some Ali-Silvey distance measures.The performance degradation of distributed detection systems relative to the cen lized system was also determined.It was also shown that the perfcmance of the decentralized system approaches the performance of the centralized system very quickly as a function of the number of quantization levels.

CHAPTER 5

PERFORMANCE EVALUATION OF DISTRIBUTED BAYESIAN DETECTION STRUCTURES

5.1 Introduction

We have previously discussed the computational difficulties associated with the procedures for the design of optimum decentralized detection systems. Therefore, the study of decentralized detection systems has been limited to small networks and very few topologies. A computationally simpler approach for the design of decentralized Bayesian detection systems was presented in Chapter 4. This approach is based upon the alternative expression for the minimum average cost derived in Chapter 2. The main objective of this chapter is to apply this approach to the design and study of four decentralized detection topologies. It is demonstrated by means of illustrative examples that relatively large networks can be handled rather easily. For these systems, we show that the design of the optimum system can be reduced to the optimization of a single function of a certain number of variables that depend upon the configuration considered. In Section 5.2, we briefly formulate the problem. In Section 5.3, we revisit the parallel fusion network with n local detectors discussed in Chapter 4. Here we deal with the more general situation where the observations received by the local detectors are not necessarily identical and where the local thresholds are not assumed to be identical. Under these conditions, we show that the design of the optimum system reduces to the optimization of a single function of n variables. This optimization is performed only once with no need to search over all the possible fusion rules, i.e., the optimization procedure is not exhuastive. In Section 5.4, we consider a variation of the parallel system where we allow the fusion center to make its judgement based on the received local decisions as well as its directly received observation (side information). While for the parallel fusion system there, the number of possible fusion rules grows exponentially with n (see Table 1.1), there are 2^n decision thresholds

for the system with side information. These decision thresholds span the entire observation space of the local observations. For this structure we show again, that the design of the optimal system reduces to the optimization of a single function of n variables. In Section 5.5, we consider the design of a hierarchical system with 2n local detectors and n regional decision makers. The regional decision makers are assumed to make their decisions based on the decisions received from the local decision makers and also on their own observations, i.e., side information. The design of the optimum system for this structure reduces to the optimization of a function of 3n variables. In Section 5.6, we discuss the design of the hierarchical system in which the regional decision makers do not have observations of their own. In Section 5.7 we present several examples along with some numerical results. Section 5.8 contains a discussion and some concluding remarks.

5.2 Problem Formulation

In Chapter 4, we developed a computationally simple approach to the design of decentralized detection systems. This procedure is based upon an alternative representation of the minimum average cost in terms of a modified form of the Kolmogorov variational distance. The procedure was applied to the design of the parallel fusion network when the incoming observations were assumed independent and identically distributed and the local thresholds were assumed to be identical. The cases of hard decisions and soft decisions were analyzed. In this Chapter we apply the design procedure to the design of a number of decentralized detection structures including the parallel fusion network. These structures are used for the binary hypothesis testing problem. The null hypothesis H_0 with a priori probability π_0 is tested against the alternative hypothesis H_1 with a priori probability π_1. The criterion we adopt is the minimization of the system average cost. We consider the case when the local as well as the regional decisions are either 0 (corresponding to H_0) or 1 (corresponding to H_1), i.e., we only treat the case of hard decisions. Here, we assume that the incoming observations are independent but not necessarily identical . Let C_{ij} ; i, j = 0, 1, be the overall cost of deciding H_i when H_j is true. Since the final decision is made at the fusion center, these

costs are applicable there, and the fusion center can be looked upon as a minimum average cost receiver. Therefore, the minimum average cost at the fusion center can be represented as given in (2.21). Now we apply our computationally simple approach to the design of several network topologies.

5.3 Parallel Fusion Network

Let us consider the system S_1 shown in Fig. 1.1 consisting of n local detectors and a global decision maker. The system receives n observations $X_1,...,X_n$ in which X_i denotes the observation received by the local detector LD_i, $1 \leq i \leq n$. We assume that these observations are independent with conditional pdf's $p_0(x_i)$ and $p_1(x_i)$ under hypotheses H_0 and H_1 respectively. Due to the bandwidth constraints on the channels linking the local detectors to the global decision maker, the local detector LD_i compresses its raw observation X_i to a single hard decision, z_i, indicating whether H_0 or H_1 is true and transmits it to the fusion center. The design procedure employed here can be easily extended to the design of decentralized detection systems employing soft decisions, but this case will not be considered here. We assume that the decision made by the local detector LD_i is independent of the decisions made by the other local detectors. Due to the independence assumption, in the optimum system, each local detector LD_i performs a local likelihood ratio test with respect to some local threshold τ_i. That is, the decisions z_i are made based on the following rule

$$z_i = \begin{cases} 1 & \text{if } p_1(x_i)/p_0(x_i) \geq \tau_i \\ \\ 0 & \text{otherwise} \end{cases} \qquad i=1,...,n \qquad (5.1)$$

where $z_i = 1$ means that H_1 has been declared true and $z_i = 0$ means that H_0 has been declared true by the local detector LD_i. The decision rule (5.1) thus characterizes each local detector LD_i by a local probability of detection PDL_i and a local probability of false alarm PFL_i. Each element of the local decision vector $\underline{U} = [z_1...z_n]$ is a Bernoulli random variable having the following density functions under the two hypothesis

$$f_0(z_i) = PFL_i^{z_i}(1-PFL_i)^{1-z_i} \qquad (5.2)$$

$$f_1(z_i) = PDL_i^{z_i}(1-PDL_i)^{1-z_i} \qquad (5.3)$$

It is to be noted that unlike the case considered in Chapter 4 where the decisions z_1, z_2, ..., z_n were independent and identically distributed, the decisions z_1, z_2, ..., z_n here are independent but are not assumed to be identically distributed. This is a result of our earlier assumption that the observations X_1, X_2, ..., X_n are not necessarily identically distributed. Using the assumption of independence between the local decisions z_i, the decision vector $\underline{U}$ has the following density function under the two hypotheses

$$f_j(\underline{U}) = \prod_{i=1}^{n} f_j(z_i), j = 0, 1 \qquad (5.4)$$

These decisions are sent over bandlimited channels to the global decision maker. The global decision u_0 is obtained by the fusion center based on the vector $\underline{U}$ by performing the likelihood ratio test (4.2) with $\Lambda(\underline{U}) = f_1(\underline{U})/f_0(\underline{U})$. Using (2.21), we can find the MAC of the system as

$$R = R_0 - \frac{1}{2}\sum_{\underline{U}} |C_1 f_1(\underline{U}) - C_0 f_0(\underline{U})| \qquad (5.5)$$

We should point out that of all the possible fusion rules, (4.2) specifies the fusion rule that achieves the smallest MAC for a given set of local decision rules. The cost corresponding to this (best) fusion rule is specified by equation (5.5). We also emphasize that an explicit knowledge of the fusion rule is not required to determine the MAC given in (5.5). This point was illustrated by means of an example in Section 4.4.

While (5.5) determines the MAC for given a priori probabilities, cost assignments, and local decision rules, it can also be used for designing the optimum system. In order to use (5.5) for this objective we let $\underline{\tau} = [\tau_1...\tau_n]$ denote the vector whose elements are the local thresholds, where τ_i is the threshold of local detector LD_i as defined in (5.1). The local detection probability PFL_i and the local false alarm probability PFL_i of local detector LD_i are functions of τ_i. We stress this by expressing PDL_i as $PDL_i(\tau_i)$ and PFL_i as $PFL_i(\tau_i)$. Therefore, we can express the conditional distributions given by (5.2) and (5.3) of the lo-

cal decisions z_i in terms of τ_i as

$$f_0(z_i(\tau_i)) = [PFL_i(\tau_i)]^{z_i}[1-PFL_i(\tau_i)]^{1-z_i} \qquad (5.6)$$

$$f_1(z_i(\tau_i)) = [PDL_i(\tau_i)]^{z_i}[1-PDL_i(\tau_i)]^{1-z_i} \qquad (5.7)$$

The conditional densities of the decision vector $\underline{U}$ correspondingly can be expressed in terms of $\underline{\tau}$ as

$$f_j(\underline{U}(\underline{\tau})) = \prod_{i=1}^{n} f_j(z_i(\tau_i)), \, j = 0, 1 \qquad (5.8)$$

For a given vector $\underline{\tau}$ of local thresholds, the observation space of $\underline{U}$ is optimally partitioned using (4.2) with $\Lambda(\underline{U})$ expressed in terms of the densities given in (5.8) as

$$\Lambda(\underline{U}(\underline{\tau})) = \frac{f_1(\underline{U}(\underline{\tau}))}{f_0(\underline{U}(\underline{\tau}))} \qquad (5.9)$$

As pointed out in the second remark in Section 4.2, the decision rule (4.2) with $\Lambda(\underline{\tau})$ as given in (5.9) specifies the fusion rule with the smallest MAC for the given vector $\underline{\tau}$. The MAC corresponding to this fusion rule with n incoming decisions expressed in terms of $\underline{\tau}$ is

$$R_n(\underline{\tau}) = R_0 - 0.5\sum_{\underline{U}}|C_1 f_1(\underline{U}(\underline{\tau})) - C_0 f_0(\underline{U}(\underline{\tau}))| \qquad (5.10)$$

where the summation is taken over all the possible values of $\underline{U}$. Equation (5.10) is a function of n variables, namely, the n local thresholds $\tau_1, ..., \tau_n$. For a given value of $\underline{\tau}$, Equation (5.10) determines the cost corresponding to the best fusion rule among all the possible fusion rules. Therefore, to optimally design the overall system, $R_n(\underline{\tau})$ should be minimized with respect to $\underline{\tau}$. The resulting fusion rule can be determined from (5.2) (see also Appendix A).

If the observations $X_1, ..., X_n$ are independent and identically distributed under both hypotheses, then the receiver operating characteristics of the n local detectors are identi-

104

cal. That is, if α_1 and β_1 are the false alarm and detection probabilities that are achieved at threshold τ for one of the local detectors, then the false alarm and detection probabilities of all the other detectors at threshold τ will also be α_1 and β_1. If we assume that these local detectors when operating in a decentralized system have identical thresholds, then we have $PDL_i = PDL$ and $PFL_i = PFL$ for $1 \le i \le n$. Eventhough this assumption is intuitively appealing, a number of counterexamples have been reported in the literature [20,58] whereby the overall system cost is minimized by nonidentical local decision rules eventhough the local detectors are identical. It has been observed in [24] that for decentralized Beyesian detection systems, the identical local decision rules assumption often results in little or no loss of optimality. The system design procedure greatly simplifies for the identical detector case. In this case the vector $\underline{\tau}$ reduces to a scalar τ and the n-variable design equation (5.10), reduces to equation (4.47) which is a function of one variable only. In Example 5.1, we will use both equations (5.10) and(5.47) to design a system consisting of six local detectors with identical observation statistics.

5.4 Parallel Fusion Network with Side Information

In this section we consider the distributed detection system S_2 shown in Fig. 1.2. This system is different from the one treated in the previous subsection in that the global decision maker receives a local observation of its own (or side information) in addition to the local decisions z_i. The observation vector based on which the fusion center makes the final decision is the augmented vector $[\underline{U}X_0]$ of the local decisions z_i and the observation X_0 at the fusion center. The global decision is made according to the rule

$$\Lambda(x_0) \frac{f_1(\underline{U})}{f_0(\underline{U})} \overset{\overset{H_1}{>}}{\underset{\underset{H_0}{<}}{}} \eta \tag{5.11}$$

where

$$\Lambda(x_0) = \frac{p_1(x_0)}{p_0(x_0)}$$

and η is the global threshold defined in equation (4.2). An equivalent test can be obtained in terms of the local observation X_0 as

$$\Lambda(x_0) \quad \begin{matrix} H_1 \\ > \\ \eta \\ < \\ H_0 \end{matrix} \quad \frac{f_0(\underline{U})}{f_1(\underline{U})} = \eta_Q, \ Q = 1,\ldots,2^n \qquad (5.12)$$

By using the expressions given in (5.4) for $f_0(\underline{U})$ and $f_1(\underline{U})$ and taking the logarithm of both sides of (5.12) we get the following (see Appendix A for details)

$$\log \eta \frac{f_0(\underline{U})}{f_1(\underline{U})} = \log \eta + \sum_{i=1}^{n} \log \frac{1 - PFL_i}{1 - PDL_i} + \sum_{i=1}^{n} z_i \log \frac{PFL_i(1 - PDL_i)}{PDL_i(1 - PFL_i)} \qquad (5.13)$$

Note that the equivalent test in terms of the logarithm specified by (5.12) and (5.13) is an extension of the optimum data fusion algorithm developed in [9]. Here, the locally received observation X_0 has been taken into account. As can be seen from this new equation (giving the test), there are 2^n different decision thresholds to be employed at the global decision maker. Each threshold η_Q, $Q = 1, \ldots, 2^n$ corresponds to a particular sequence of the 2^n possible values that the decision vector $\underline{U}$ takes. Using (2.20) we can determine the minimum cost of the system as

$$R = R_0 - \frac{1}{2} \int \sum_{\underline{U}} |C_1 f_1(\underline{U}) p_1(x_0) - C_0 f_0(\underline{U}) p_0(x_0)| dx_0 \qquad (5.14)$$

where the integration is performed over all values of X_0. When the observations $X_1 \ldots X_n$ are independent and identically distributed under both hypotheses, the receiver operating characteristics of the n local detectors become identical. If we assume, as we did in the previous subsection, that the local detectors use identical thresholds, then the equivalent thresholds given in (5.13) simplify to

$$\log \eta \frac{f_0(\underline{U})}{f_1(\underline{U})} = \log \eta \left(\frac{1 - PFL}{1 - PDL}\right)^n + \log \frac{PFL(1 - PDL)}{PDL(1 - PFL)} \sum_{i=1}^{n} z_i \qquad (5.15)$$

Equation (5.15) indicates that there are (n+1) decision thresholds employed at the global

decision maker. If we let K be the number of local thresholds that decide in favor of H_1, then K has the distributions given by (4.34) and (4.35). Using (4.34) and (4.35) the MAC in (5.14) can be expressed as

$$\mathfrak{R} = R_0 - \frac{1}{2} \int \sum_{k=0}^{n} \binom{n}{k} \left| C_1 PDL^k (1 - PDL)^{n-k} p_1(x_0) - C_0 PFL^k (1 - PFL)^{n-k} p_0(x_0) \right| dx$$

(5.16)

As before, we let $\underline{\tau} = [\tau_1 ... \tau_n]$ be the vector of local thresholds. We express PDL_i as $PDL_i(\tau_i)$ and PFL_i as $PFL_i(\tau_i)$. For any setting of the local thresholds $\underline{\tau}$, the optimum partitioning of the observation space of X_0 is made according to the decision rule (5.12). The decision thresholds η_Q expressed explicitly in terms of $\underline{\tau}$ are

$$\eta_Q(\underline{\tau}) = \eta \frac{\displaystyle\prod_{i=1}^{n} f_0(z_i(\tau_i))}{\displaystyle\prod_{i=1}^{n} f_1(z_i(\tau_i))}$$

(5.17)

where the density functions in the numerator and the denominator are those given in equations (5.6) and (5.7). The resulting system MAC from (5.14) can be expressed as a function of $\underline{\tau}$ as

$$R_n(\underline{\tau}) = R_0 - \frac{1}{2} \int \left(\sum_{\underline{U}} \left| C_1 f_1(\underline{U}(\underline{\tau})) p_1(x_0) - C_0 f_0(\underline{U}(\underline{\tau})) p_0(x_0) \right| \right) dx_0$$

(5.18)

If the local threshold vector $\underline{\tau}$ is changed to $\underline{\tau}''$ it results in a different set of 2^n decision thresholds at the fusion center and, consequently, a different value of the global MAC, $R_n(\underline{\tau}'')$. Our goal is to obtain the best global MAC. Therefore, $R_n(\underline{\tau})$ should be minimized with respect to the local threshold vector $\underline{\tau}$. The resulting system, i.e., the resulting local threshold vector $\underline{\tau}$ and the corresponding 2^n decision thresholds specified by (5.12), is optimum. It should be emphasized that equation (5.18) is a function of n variables, namely, the n local thresholds. The explicit dependence of the cost on the 2^n decision thresholds at the fusion center has been avoided.

The design equation (5.18) simplifies tremendously for the identical thresholds case. In this case the vector $\underline{\tau}$ reduces to a scalar threshold τ. The decision thresholds specified by (5.17) become

$$\eta_k = \eta \, \frac{PFL^k(\tau)[1-PFL(\tau)]^{n-k}}{PDL^k(\tau)[1-PDL(\tau)]^{n-k}} \quad ,k = 0, ..., n \quad\quad (5.19)$$

The system MAC in terms of τ, $R_n(\tau)$, is obtained by explicitly expressing *PFL* as *PFL*(τ) and *PDL* as *PDL*(τ) in equation (5.16). In this case the cost function is a function of one variable, namely, the common local threshold. The global minimum cost is obtained by minimizing $R_n(\tau)$ with respect to the local threshold τ. The resulting local threshold along with the corresponding (n+1) decision thresholds determined by (5.19) specify the optimum system. An example will be presented in Section 5.7.

5.5 Hierarchical System with Side Information at the Regional Detectors

In this subsection we consider the hierarchical decentralized detection system S_3 shown in Fig. 1.3. The system consists of 2n local decision makers, n regional decision makers (RD's) and a global decision maker. Local detectors LD_{2i-1} and LD_{2i} process their locally received observations X_{2i-1} and X_{2i} and forward their decisions z_{2i-1} and z_{2i} to an intermediate regional decision maker RD_i, i = 1, ..., n. In the hierarchical system considered in this chapter, the local decisions of only two detectors are combined at the regional decision makers. The results can be generalized to the case of more than two local detectors per regional decision maker. However, for simplicity in presentation, we consider the case of two local detectors only. The regional detector RD_i combines the two local decisions along with its directly received observation Y_i to make the regional decision u_i. The observations Y_j, j = 1, ..., n are assumed to be independent with pdf's $p_0(y_j)$ and $p_1(y_j)$ under the hypotheses H_0 and H_1 respectively. The decision vector $\underline{U} = [u_1...u_n]$ is used by the global decision maker to make the final decision u_0. The local decisions z_i, i = 1, 2, ..., 2n are made based on the decision rule (5.1) and hence, each local detector is characterized by a local probability of detection PDL_i and a local probability of false alarm PFL_i. The fi-

108

nal decision is made based on the decision rule (4.2) with $\Lambda(\underline{U})$ to be specified later. Let η_i, $i = 1, ..., n$ be the unknown threshold of the regional detector RD_i. The regional decision u_i is made according to the test

$$\frac{p_1(y_i)f_1(z_{2i-1})f_1(z_{2i})}{p_0(y_i)f_0(z_{2i-1})f_0(z_{2i})} \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \eta_i \ , i = 1, 2, ..., n \qquad (5.20)$$

where $f_j(z_i)$; $j = 0, 1$, $i = 1, 2, ..., 2n$ are given by equations (5.2) and (5.3). The test (5.20) can also be expressed in terms of the regional observation Y_i as follows

$$\frac{p_1(y_i)}{p_0(y_i)} \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \eta_i \frac{f_0(z_{2i-1})f_0(z_{2i})}{f_1(z_{2i-1})f_1(z_{2i})} = \eta_{ikl} \ , k, l = 0, 1 \qquad (5.21)$$

where k is equal to z_{2i-1}, the decision of local detector LD_{2i-1}, and l is equal to z_{2i}, the decision of local detector LD_{2i}. Depending on the values that z_{2i-1} and z_{2i} take, the regional decision maker RD_i employs one of the following thresholds

$$\eta_{i00} = \eta_i \frac{(1 - PFL_{2i-1}) \ (1 - PFL_{2i})}{(1 - PDL_{2i-1}) \ (1 - PDL_{2i})} \qquad (5.22\text{-a})$$

$$\eta_{i01} = \eta_i \frac{(1 - PFL_{2i-1}) \ PFL_{2i}}{(1 - PDL_{2i-1}) \ PDL_{2i}} \qquad (5.22\text{-b})$$

$$\eta_{i10} = \eta_i \frac{PFL_{2i-1} \ (1 - PFL_{2i})}{PDL_{2i-1} \ (1 - PDL_{2i})} \qquad (5.22\text{-c})$$

$$\eta_{i11} = \eta_i \frac{PFL_{2i-1}PFL_{2i}}{PDL_{2i-1}PDL_{2i}} \qquad (5.22\text{-d})$$

Let $PDR_i(\eta_{ikl})$ and $PFR_i(\eta_{ikl})$ denote detection and false alarm probabilities of regional decision maker RD_i for a given threshold η_{ikl}, i.e., when $z_{2i-1} = k$ and $z_{2i} = l$. These probabilities are given as

$$PDR_i(\eta_{ikl}) = P\ (u_i = 1 | z_{2i-1} = k, z_{2i} = l, H_1) \qquad (5.23\text{-a})$$

$$PFR_i(\eta_{ikl}) = P\ (u_i = 1 | z_{2i-1} = k, z_{2i} = l, H_0) \qquad (5.\ 23\text{-b})$$

Then using the theorem of total probability, we can find the unconditional detection and false alarm probabilities as

$$PDR_i = \sum_k \sum_l P(z_{2i-1} = k,\ z_{2i} = l\ | H_1)\ PDR_i(\eta_{ikl}) \qquad (5.24\text{-a})$$

$$PFR_i = \sum_k \sum_l P(z_{2i-1} = k,\ z_{2i} = l\ | H_0)\ PFR_i(\eta_{ikl}) \qquad (5.24\text{-b})$$

Expanding over k and l we can express PDR_i and PFR_i as

$$PDR_i = (1\text{-}PDL_{2i-1})(1\text{-}PDL_{2i})PDR_i(\eta_{i00}) + (1\text{-}PDL_{2i-1})PDL_{2i}\ PDR_i(\eta_{i01})$$

$$+PDL_{2i-1}(1\text{-}PDL_{2i})PDR_i(\eta_{i10}) + PDL_{2i-1}PDL_{2i}\ PDR_i(\eta_{i11}) \qquad (5.25\text{-a})$$

$$PFR_i = (1\text{-}PFL_{2i-1})(1\text{-}PFL_{2i})PFR_i(\eta_{i00}) + (1\text{-}PFL_{2i-1})PFL_{2i}\ PFR_i(\eta_{i01})$$

$$+PFL_{2i-1}(1\text{-}PFL_{2i})PFR_i(\eta_{i10}) + PFL_{2i-1}PFL_{2i}\ PFR_i(\eta_{i11}) \qquad (5.25\text{-b})$$

Each regional detector RD_i is characterized by a detection probability PDR_i and a false

alarm probability PFR_i. The density functions of the regional decision u_i under H_0 and H_1 are given by the following Bernoulli distributions

$$f_0(u_i) = PFR_i^{u_i}(1-PFR_i)^{1-u_i} \qquad (5.26\text{-b})$$

$$f_1(u_i) = PDR_i^{u_i}(1-PDR_i)^{1-u_i} \qquad (5.26\text{-a})$$

for $u_i = 0, 1, i = 1, ..., n$. Using the independence assumption between the regional decisions, we obtain the following density functions for the decision vector

$$f_j(\underline{U}) = \prod_{i=1}^{n} f_j(u_i), j = 0, 1 \qquad (5.27)$$

The system MAC is computed using (5.5) with $f_0(\underline{U})$ and $f_1(\underline{U})$ as given in (5.27). Let $\underline{\tau} = [\tau_1...\tau_{2n}]$ be the vector of local thresholds and $\underline{\eta} = [\eta_1...\eta_n]$ be the vector of regional thresholds. The design of the optimum system calls for the determination of $\underline{\tau}$, $\underline{\eta}$, and the fusion rule that minimize the global cost. For a given local threshold τ_j, $j = 1, 2, ..., 2n$, the local detection and false alarm probabilities PDL_j and PFL_j are functions of τ_j as discussed in the previous sections. For a given regional threshold η_i, $i = 1, ..., n$, the conditional regional thresholds given in (5.22) are functions of the local thresholds τ_{2i-1} and τ_{2i} as evident from the the dependence of the thresholds in (5.22) on the local detection and false alarm probabilities. These thresholds are also functions of η_i as can be seen from (5.21). Consequently, the unconditional regional detection and false alarm probabilities given in (5.25) are functions of the variables τ_{2i-1}, τ_{2i}, and η_i. More explicitly, we write PDR_i as $PDR_i(\tau_{2i-1}, \tau_{2i}, \eta_i)$ and PFR_i as $PFR_i(\tau_{2i-1}, \tau_{2i}, \eta_i)$. In terms of the system variables, the system MAC in (5.5) becomes

$$R(\underline{\tau}, \underline{\eta}) = R_0 - \frac{1}{2}\sum_{\underline{U}} \left| C_1 \prod_{i=1}^{n} f_1(u_i(\eta_i, \tau_{2i-1}, \tau_{2i})) - C_0 \prod_{i=1}^{n} f_0(u_i(\eta_i, \tau_{2i-1}, \tau_{2i})) \right|$$

$$(5.28)$$

Equation (5.28) is a function of 3n variables. In order to obtain the optimum system, equa-

tion (5.28) has to be minimized with respect to these variables. The fusion rule can be obtained from (4.2) once the optimum local and regional thresholds are determined. With the identical threshold assumption the design equation (5.28) becomes a function of two variables, namely, the common local threshold τ and the common regional threshold η.

It should be pointed out that the performance of this system is expected to be inferior to those in the preceding sections for the same number of observations. The reason for this is the extra data compression at the regional level, so that the information available to the fusion center is less than before. This will be illustrated in Section 5.7 where several numerical examples are considered.

5.6 Hierarchical Decentralized Detection System

In this subsection we consider the hierarchical detection system S_4 shown in Fig. 1.4. The difference between this system and the one treated in the previous section is that the regional decision makers have no observations of their own. Therefore, they have to make their decisions solely on the basis of the local decisions they receive. The local decisions and the final decisions are still made as in Section 5.5, i.e., the local decisions are made on the basis of the local observations and the global decision is made on the basis of the decisions received from the regional decision makers. However, due to the unavailability of observations at the regional detectors, special attention has to be paid to decision making at the regional decision makers. In the rest of this section, we modify the optimization procedure of Section 5.5 to take into account the absence of side information at the regional detectors.

Let η_i be the threshold used by regional decision maker RD_i. The regional decision u_i is made based on the test

$$L(z_{2i-1}, z_{2i}) = \frac{f_1(z_{2i-1})f_1(z_{2i})}{f_0(z_{2i-1})f_0(z_{2i})} \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \eta_i \qquad (5.29)$$

112

Substituting the density functions for z_k given by (5.2) and (5.3) into (5.29) we get

$$\frac{(PDL_{2i-1})^{z_{2i-1}} (1-PDL_{2i-1})^{1-z_{2i-1}} (PDL_{2i})^{z_{2i}} (1-PDL_{2i})^{1-z_{2i}}}{(PFL_{2i-1})^{z_{2i-1}} (1-PFL_{2i-1})^{1-z_{2i-1}} (PFL_{2i})^{z_{2i}} (1-PFL_{2i})^{1-z_{2i}}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta_i \quad (5.30)$$

Taking the logarithm of both sides of (5.30) and arranging terms we obtain the following

test

$$A_{2i-1} z_{2i-1} + A_{2i} z_{2i} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} C_i \quad (5.31)$$

where

$$A_{2i-1} = \log \frac{PDL_{2i-1} (1 - PFL_{2i-1})}{PFL_{2i-1} (1 - PDL_{2i-1})}$$

$$A_{2i} = \log \frac{PDL_{2i} (1 - PFL_{2i})}{PFL_{2i} (1 - PDL_{2i})}$$

$$C_i = \log \eta_i \frac{(1 - PFL_{2i-1}) (1 - PFL_{2i})}{(1 - PDL_{2i-1}) (1 - PDL_{2i})}$$

The random variables z_{2i-1} and z_{2i} are Bernoulli random variables under both hypotheses, but their linear combination in (5.31) is not. Let

$$L_i = A_{2i-1} z_{2i-1} + A_{2i} z_2 \quad (5.32)$$

113

be the sufficient statistic at the regional decision maker RD_i. The regional probabilities of false alarm and detection are

$$PFR_i = P \{u_i = 1 | H_0 \text{ is } true\} = P \{L_i \geq C_i | H_0\} \qquad (5.33\text{-a})$$

$$PDR_i = P \{u_i = 1 | H_1 \text{ is } true\} = P \{L_i \geq C_i | H_1\} \qquad (5.33\text{-b})$$

Now we show how equation (5.28) can be used to design the system. In Section 5.5, a likelihood ratio test at the regional detectors was formulated where the likelihood ratio was obtained using y_i as the observations and the incoming local decisions were used to modify the threshold η_i. In that case, the receiver operating characteristic at the regional decision maker was continuous and all values of η_i were permissible. Thus it was possible to express the regional probabilities of detection and false alarm in terms of the regional threshold η_i and the probabilities of detection and false alarm of the local detectors connected to this regional detector. The optimum system was obtained by optimizing the cost in (5.28) with respect to the local and regional thresholds. All of these thresholds were assumed to be independent, i.e., the number of independent variables was 3n. The difficulty in the design of the system of Fig. 1.4 arises due to the fact that the regional thresholds and the local thresholds cannot be assumed to be independent, i.e., we do not have 3n independent variables. In Appendix C, we show that for given local thresholds τ_{2i-1} and τ_{2i}, the regional threshold η_i may lie in one of five possible regions. These regions are determined in terms of the local thresholds. If the regional threshold η_i lies in the two outside regions, it results in maximum possible values of system cost. Therefore, these two regions are not desirable. In the remaining three regions the system cost does not change as the regional threshold η_i is varied within any of these regions. Thus it is not possible to express the regional probabilities of detection and false alarm in terms of the regional threshold η_i explicitly. The functional relationship between τ_{2i-1} and τ_{2i} and η_i can be further emphasized by expressing η_i as $\eta_i(\tau_{2i-1}, \tau_{2i})$. The optimization procedure of Section 5.5 can now be used except that we have only two independent variables τ_{2i-1} and τ_{2i} for each regional detector RD_i.

5.7. Examples

In this section we present several numerical examples that illustrate the utility of the design procedure described in this chapter. Our goals in presenting these examples are

1) To examine the effect on the global MAC of allowing identical local detectors to use identical local decision rules.

2) To examine the effect of having side information at the regional and global decision makers.

3) To compare the MPOE of the various systems when the total number of observation received by each system is six.

In these examples we assume that under hypothesis H_0, each local and regional observation is a Gaussian random variable with mean zero and variance σ_i^2. While under hypothesis H_1, each observation is a Gaussian random variable with mean μ_i and variance σ_i^2. As discussed in Example 5.2, for the Gaussian hypothesis testing problem, it is more convenient to express the local thresholds in terms of the local observations as (see equation (4.71)).

$$\Gamma_i = \frac{\sigma_i^2}{\mu_i} \log \tau_i + \frac{\mu_i}{2} \tag{5.34}$$

where τ_i represent the local thresholds defined in equation (5.1). Note that the thresholds that appear in the tables at the end of this section are those given by (5.34). We should mention that in all the systems, the design procedure reduces to the minimization of a cost function. In performing the optimization we use the method developed by Hooke and Jeeve [69].

Example 5.1

The objective of this example is to examine whether optimality is lost by assuming that the identical local detectors in systems S_1 and S_2 are using the same threshold.

Consider the system S_1 with six local detectors and the system S_2 with five local de-

tectors so that both the systems have an identical number of local observations, i.e., six local observations. Let $\pi_0 = 0.5$, $\mu_i = 2.5$, and $\sigma_i^2 = 1$ for all the observations in the two systems. Using different cost assignments we perform the design of the systems for two cases. In the first one, we do not restrict the local detectors to employ an identical threshold, i.e., in obtaining the optimum local thresholds, we assume that each one of the local detectors employs a different threshold. The appropriate design equations in this case are (5.10) for system S_1 and (6.18) for system S_2 which are functions of six variables and five variables respectively. In the second case, we assume that all local detectors have identical thresholds. The appropriate design equations are (4.47) for system S_1 and (5.18) for system S_2. For both systems, the thresholds obtained with and without equal threshold restriction were identical for a variety of cost assignments. In other words, for this example optimality is not lost if identical threshold assumption is made when observation statistics at the sensors are identical. The values of the optimum local thresholds and the optimum costs are shown in Tables 5.1 and 5.2. Also shown in Table 5.1 are the optimum fusion rules for system S_1. These fusion rules are obtained by substituting the optimum local thresholds in the likelihood ratio function (5.9) and using (4.2) (see also Appendix A). In Table 5.2, we have also included the cost for the optimum centralized system, i.e., where all six observations are processed centrally. The performance loss due to decentralization is evident from Table 5.2. Also, note that decentralized detection system with side information performs better than the system without side information.

Example 5.2

In this example, we examine the effect of havind one noisy local detector on the performance of the decentralized systems S_1 and S_2.

Consider the hypothesis testing problem of Example 5.1 for the MPOE criterion. We assume that the assumptions and conditions of that example are still valid except for one of the local detectors where we assume that its noise variance σ^2 is a variable. For S_1, this means that for five of the local detectors the observation statistics under the two hypotheses are given as

C_{00}	C_{11}	C_{10}	C_{01}	Optimum local threshold	Optimum cost	Optimum fusion rule
0	0	1	1	1.46	6.059×10^{-3}	3 out of 6
0	0	1	5	1.30	1.266×10^{-2}	3 out of 6
0	0	5	1	1.20	1.266×10^{-2}	4 out of 6
1	2	4	6	1.43	1.5208	4 out of 6
0	0	30	40	1.01	0.2107	4 out of 6
1	1	2	2	1.037	1.006	4 out of 6

Table 5.1: Results of Example 5.1 for the parallel fusion network S_1

C_{00}	C_{11}	C_{10}	C_{01}	Optimum local threshold	Optimum cost for system S_2	Optimum cost for the centralized system
0	0	1	1	1.24	5.216×10^{-3}	1.0998×10^{-3}
0	0	1	5	1.15	1.119×10^{-2}	2.3877×10^{-3}
0	0	5	1	1.36	1.1192×10^{-2}	2.3877×10^{-3}
1	2	4	6	1.23	1.518	1.5038
0	0	30	40	1.23	0.18046	0.03806
1	1	2	2	1.247	1.00522	1.0011

Table 5.2 : Results of Example 5.1 for the parallel network S_2 with side information

$$H_0 : X_i \sim N(0,1)$$

$$H_1 : X_i \sim N(2.5,1) \, , \, i = 1,2,...,5.$$

while for the sixth local detector the observation statistic is given as

$$H_0 : X_6 \sim N(0,\sigma^2)$$

$$H_1 : X_6 \sim N(2.5,\sigma^2)$$

For S_2, the observation statistics under the two hypotheses for four of the local detectors and the side observation are given as

$$H_0 : X_i \sim N(0,1)$$

$$H_1 : X_i \sim N(2.5,1) \, , \, 0 = 1, 2, ...,4.$$

while for the fifth local detector the observation statistic is given as

$$H_0 : X_5 \sim N(0,\sigma^2)$$

$$H_1 : X_5 \sim N(2.5,\sigma^2)$$

The objective is to examine the performance of the systems as a function of the noise level σ^2 of one of the detectors when the rest of the system remains fixed. The cost of each system is obtained by minimizing the appropriate system cost equation. The solid curves in Figures 5.1 and 5.2 respectively show the MPOE for systems S_1 and S_2 versus σ^2. The broken line in Figure 5.1, shows the MPOE of the parallel fusion system when the noisy detector is disregarded, i.e., the MPOE when system makes the final decision based on the five non-noisy detectors. Similarly, the broken line in Figure 5.2 represents the MPOE of the parallel fusion system with side information when the final decision is made based on four local detectors (excluding the noisy detector) and the side observation. The broken lines in Figures 5.1 and 5.2 are constants for all the values of σ^2 because the decision of the noisy detector is ignored while making the decision.

Example 5.3

In this example we compare the optimum performance of system S_1 with n local detectors to the optimum performance of system S_2 with (n-1) local detectors for various
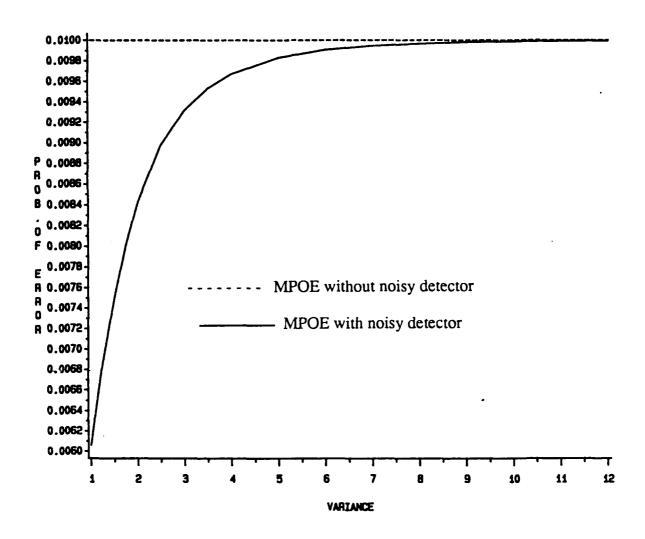
Fig. 5.1. MPOE of system S_1 when the variance of the noise at the noisy detector is varied.
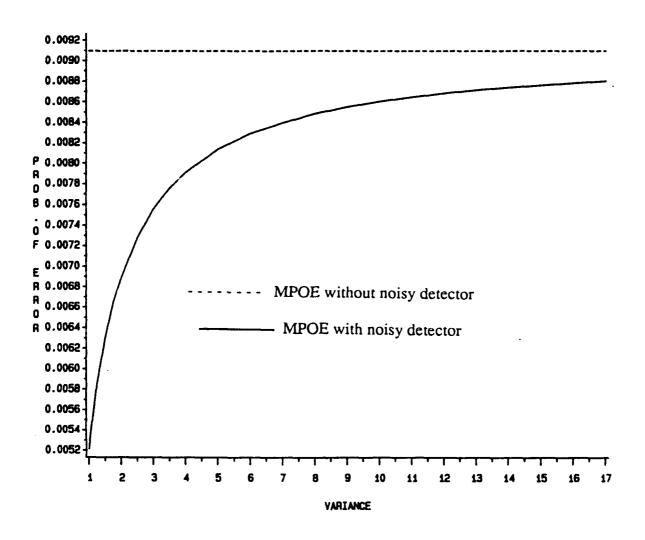
Fig. 5.2. MPOE of system S_2 when the variance of the noise at the noisy detector is varied

values of n.

Let $\mu_i = 1$ and $\sigma_i^2 = 1$ for all the observations in the two systems. Also let $C_{00} = C_{11} = 0$ and $C_{10} = C_{01} = 1$. First, let us compare the performance of system S_1 to the performance of system S_2 for the case when $n = 2$ (see Example 4.2). Note that when $n = 2$, the system S_2 reduces to the two-stage serial system. The comparison is made for all values of π_0; $0 \leq \pi_0 \leq 1$. The optimum minimum probability of error (MPOE) of system S_1 and S_2 as a function of π_0 is shown in Fig. 5.3. It is clear from this figure that S_2 performs better than S_1 for all values of π_0. These results are consistent with the results obtained in [21, 23]. In Figure 5.3, we also present the performance of the optimum centralized system. At the point $\pi_0 = 0.5$, we see that the MPOE of system S_1 is 12.9 % more than that of the centralized system. For system S_2 the corresponding percentage is 7.4 % only. Next, we compare the performance of systems S_1 and S_2 as a function of n for the case when $\pi_0 = 0.5$. Based on the observation of the results of Example 5.1, here we design both systems under the identical thresholds assumption without loss of optimality. For each value of n, the systems S_1 and S_2 are designed using (4.48) and (5.16) respectively. The optimum thresholds and the optimum MPOE are shown in Table 5.3 along with the MPOE of the centralized system with n observations. As can be seen from Table 5.3, the MPOE of the parallel system with 2 local detectors is 5.2 % higher than the MPOE of the corresponding serial system. For other values of n this percentage is smaller. As expected, for large n both systems have essentially the same performance. As n gets larger, the percentage by which the MPOE of both systems are higher than that of the centralized system becomes larger. For the parallel system with $n = 2$, this percentage as mentioned above is 12.9 %, while for the case with $n = 10$ it is 73.6 %. Thus the need for having soft decisions rather than hard decisions becomes more apparent for larger n if the overall MPOE is to be within some reasonable bounds of the MPOE of the centralized system. This requires the transmission of more bits of information to the fusion center. This problem of soft decisions was treated in Chapter 4.
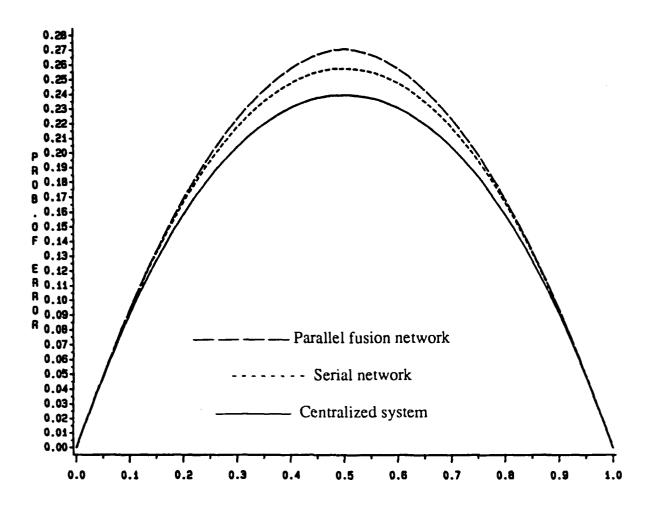
Fig. 5.3. Performance comparison for the parallel fusion network S_1 and the serial network S_2 for n = 2.

n	S_1 with n local detectors		S_2 with (n-1) local detectors		Centralized
	Threshold	MPOE	Threshold	MPOE	MPOE
2	1	0.271	0.5	0.2575	0.2397
3	0.5	0.2268	0.5	0.22244	0.1932
4	0.782	0.2014	0.5	0.19397	0.1586
5	0.5	0.1745	0.5	0.17078	0.13176
6	0.695	0.156	0.5	0.15108	0.11032
7	0.5	0.1373	0.5	0.13434	0.09293
8	0.649	0.1234	0.5	0.1198	0.07864
9	0.5	0.1095	0.5	0.10718	0.0668
10	0.62	0.09881	0.5	0.09608	0.05692
11	0.5	0.08824	0.5	0.08634	0.04862
12	0.60	0.07979	0.5	0.07769	0.04163
13	0.5	0.07157	0.5	0.07003	0.03571
14	0.5875	0.06484	0.5	0.06319	0.03068
15	0.5	0.05835	0.5	0.05713	0.0264

Table 5.3 : Results of Example 5.3 for the comparison of the MPOE of systems S_1 and S_2 for various values of the number of local detectors n.

Example 5.4

Consider the hierarchical system S_3 consisting of four local detectors and two regional detectors. We compare the MPOE of this system to the MPOE of the hierarchical system S_4 with six local detectors and three regional detectors such that the number of observations in both systems is six. Let $\pi_0 = 0.5$, $\mu_i = 2.5$, and $\sigma_i^2 = 1$ for all the observations. The optimum system is obtained by minimizing the cost in (5.28). The minimum cost (MPOE) for system S_3 is 1.0809×10^{-2}. The local thresholds were found to be identical and equal to 1.47. The regional thresholds were also identical and we obtained $\eta_1 = \eta_2 = 10.09$. For system S_4, the minimum cost was found to be 1.1874×10^{-2}. It turns out that there are two possible solutions that achieve this MPOE. In Table 5.4, the values of the optimum local thresholds are given in addition to the fusion rules employed at the regional detectors. Also given in this table are the permissible regions of the regional threshold at the regional detectors. These regions are derived in Appendix C. It is to be noted from this table, that the optimum solutions are achieved with nonidentical local thresholds. However, the local thresholds of the detectors linked to the same regional detector were found to be identical. For the sake of completeness, we summarize below the performance of the systems S_1, S_2, S_3, and S_4 for the MPOE criterion assuming that $\mu_i = 2.5$ and $\sigma_i^2 = 1$ for all of the observation

MPOE of S_1 with 6 local detectors	6.0596×10^{-3}
MPOE of S_2 with 5 local detectors	5.2134×10^{-3}
MPOE of S_3 with 4 local detectors and 2 regional detectors	1.0809×10^{-2}
MPOE of S_4 with 6 local detectors and 3 regional detectors	1.1874×10^{-2}
MPOE of the centralized system with 6 observations	1.0998×10^{-3}

5.7 Summary and Conclusions

In this chapter we considered the design and performance evaluation of four decentralized Bayesian detection structures. The procedure employed in designing these systems is based upon an alternate representation of the minimum average cost in terms of a modified form of the Kolmogorov variational distance. For the parallel fusion networks

	First solution	Second solution
Threshold at LD_1	1.725	1.77
Threshold at LD_2	1.725	1.77
Fusion rule at RD_1	OR	OR
Threshold at LD_3	0.728	0.775
Threshold at LD_4	0.728	0.775
Fusion rule at RD_2	AND	AND
Threshold at LD_5	0.725	1.77
Threshold at LD_6	0.725	1.77
Fusion rule at RD_3	AND	OR

Permissible regions of η_1

$D_{11} = 5.23 \times 10^{-3}$	$D_{12} = 4.228$	$D_{11} = 5.9 \times 10^{-2}$	$D_{12} = 4.88$
$D_{12} = 4.228$	$D_{13} = 4.228$	$D_{12} = 4.88$	$D_{13} = 4.88$
$D_{13} = 4.228$	$D_{14} = 341.42$	$D_{13} = 4.88$	$D_{14} = 405.05$

Permissible regions of η_2

$D_{21} = 2.47 \times 10^{-3}$	$D_{22} = 0.2048$	$D_{21} = 2.9 \times 10^{-3}$	$D_{22} = 0.236$
$D_{22} = 0.2048$	$D_{23} = 0.2048$	$D_{22} = 0.236$	$D_{23} = 0.236$
$D_{23} = 0.2048$	$D_{24} = 16.97$	$D_{23} = 0.236$	$D_{24} = 19.09$

Permissible region of η_3

$D_{31} = 2.47 \times 10^{-3}$	$D_{32} = 0.2048$	$D_{31} = 5.9 \times 10^{-2}$	$D_{32} = 4.88$
$D_{32} = 0.204$	$D_{33} = 0.2048$	$D_{32} = 4.88$	$D_{32} = 4.88$
$D_{33} = 0.204$	$D_{34} = 16.97$	$D_{33} = 4.88$	$D_{34} = 405.05$

Table 5.4 : Results of Example 5.4 for the hierarchical system S_4

with and without side information at the fusion center, the design of the optimum system was shown to reduce to the optimization of a single function of n variables. This optimization is performed once and does not require an exhaustive search. The case of identical observations at the local detectors was considered in detail.It was shown that the design procedure requires the optimization of a single function of one variable. For the identical observations case, the validity of the identical decision rule (at local detectors) assumption was investigated by means of an example. For the example considered, optimality was not lost when identical decision rules were assumed. For the hierarchical systems with and without side information at the regional detectors, the design of the optimum system was shown to reduce to the optimization of a function of 3n variables. For the hierarchical system with side information in the identical observations case, the local decision rules and regional decision rules were found to be identical. However, for the hierarchical system without side information, this was not true, i.e., the decision rules associated with the same regional decision maker were identical but decision rules of local detectors associated with different regional decision makers were not identical. The examples also illustrate the anticipated result that the performance of hierarchical systems is worse than that of the parallel systems due to the extra loss of information at the regional decision makers. In this chapter, the validity of our design methodology was demonstrated by applying it to four decentralized detection structures. It can also be applied to other decentralized detection topologies and structures.

CHAPTER 6

DECENTRALIZED DETECTION SYSTEM DESIGN USING THE NEYMAN-PEARSON CRITERION

6.1 Introduction

In Chapters 4 and 5 the objective was to design globally optimum minimum average cost decentralized detection systems. A number of system topologies were considered. The design procedure required the knowledge of the a priori probabilities and the costs incurred by each course of action. When such information is not available, the Neyman-Pearson criterion may be used for system design. In this case, the probability of false alarm is restricted not to exceed a prespecified value α, and the objective is to maximize the probability of detection β. Srinivasan [15] used the Neyman-Pearson criterion to obtain the local decision rules in the parallel fusion network assuming that the fusion center is a combinational logic circuit. Hoballah and Varshney [16] treated the problem in two respects. First, when the fusion rule is known and the objective is to find the local decision rules. Second, when the decision rules at the detectors are given,and the objective is to find the optimum fusion rule. Thomopoulos, Viswanathan, and Bougoulias [17] employed the Neyman-Pearson criterion for system design where both the decisions made by the individual sensors and the global decision made by the fusion center are based on the Neyman-Pearson test. Tsitsiklis [18] addressed the question of concavity of the receiver operating characteristic of the system. He found that for a given strategy, the receiver operating characteristic is not necessarily a concave function (a numerical example is provided in [19]). However, concavity can be achieved by randomizing with respect to the possible strategies. A similar result was obtained by Willet and Warren [20]. Viswanathan, Thomopoulos, and Tumuluri [21] applied the Neyman-Pearson criterion to the design of the serial decentralized configuration. They found that for the case of two sensors, the optimal serial network has a better performance than the parallel scheme (better here refers

to higher probability of detection for the same false alarm probability). While this interesting result is true for the case of two sensors, the numerical examples provided in [21] show that this result is not true, in general, for systems with more than two sensors.

As mentioned earlier, the final decision in the parallel fusion system of Fig. 1.1 is based on the decision vector $\underline{U} = [z_1 \, z_2 \, ... \, z_n]$. The observation space corresponding to this decision vector is , of course, discrete. When the observation space in a hypothesis testing problem is discrete, questions related to the randomization of the decision rule arise. In Chapter 2 we addressed the question of randomization when the decision maker was assumed to be a minimum average cost receiver. In that chapter we found that randomization of the decision rule is not necessary. This idea was implemented in Chapters 4 and 5 to design globally optimum minimum average cost decentralized detection systems without the need to randomize the decision rule. This is not necessarily the case when the design criterion is the Neyman-Pearson criterion. In this chapter, the objective is to investigate whether randomization is necessary at the fusion center in the design of the Neyman-Pearson parallel fusion system . However, the objective is not to provide a complete design approach to the Neyman-Pearson decentralized detection system. In Section 6.2 we formulate the problem. In Section 6.3 we consider the problem of randomization in detail and show that randomization is not necessary. In Section 6.4 we discuss the results obtained in this chapter.

6.2 Problem Formulation

Consider the parallel fusion system shown in Figure 1.1 consisting of n local detectors and a fusion center. The system receives n observations $X_1, X_2, ..., X_n$ where X_i denotes the observation received by the local detector LD_i. We assume that these observations are independent and identically distributed random variables with conditional probability density functions $p_j(x_i)$, $j = 0, 1$, $i = 1, ..., n$. We make the same assumption that we made in Chapter 4 in that the local detectors are identical, each one characterized by a probability of false alarm P_F and a probability of detection P_D. The operating point (P_F, P_D) can be anywhere on the receiver operating characteristic of the local detectors.

128

Let K be the random variable representing the number of sensors that decide in favor of H_1. The conditional probability density functions of K under hypotheses H_0 and H_1 are given by (4.34) and (4.35) respectively. Let $\phi(k)$ be the probability of accepting H_1 after observing $K = k$. Then, in general, $\phi(k)$ is given by (see Appendix A equation (A.13))

$$\phi(k) = \begin{cases} 1 & \text{when } k > k^* \\ \gamma & \text{when } k = k^* \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

where γ is the randomization factor and k^* is the threshold. The randomization factor in (6.1) is introduced so as to achieve the constraint on the global false alarm probability α. The probability of false alarm α is the expected value of $\phi(k)$ when H_0 is true, while the probability of detection β is the expected value of $\phi(k)$ when H_1 is true. These quantities are therefore expressed as

$$\alpha = E_0\{\phi(k)\} \tag{6.2}$$

$$\beta = E_1\{\phi(k)\} \tag{6.3}$$

Using (4.34) and (4.35), we obtain the following expressions for α and β

$$\alpha = \gamma \binom{n}{k^*} P_F^{k^*} (1 - P_F)^{n - k^*} + \sum_{k = k^* + 1}^{n} \binom{n}{k} P_F^k (1 - P_F)^{n - k} \tag{6.4}$$

$$\beta = \gamma \binom{n}{k^*} P_D^{k^*} (1 - P_D)^{n - k^*} + \sum_{k = k^* + 1}^{n} \binom{n}{k} P_D^k (1 - P_D)^{n - k} \tag{6.5}$$

When the local decision rules are known, the randomization factor and the threshold k^* can be easily determined using (6.4) to satisfy the desired probability of false alarm. The values of k^* and γ that satisfy (6.4) are used to find the probability of detection in (6.5).

The resulting β is the maximum that can be achieved. When the local decision rules are not known, and the objective is to design the system that maximizes β for the specified α, the problem becomes involved. In this case, only the constraint on the system probability of false alarm is given while three parameters have to be determined, namely, the threshold k^*, the randomization factor γ, and the local threshold. For this problem, we will investigate the role of the randomization factor γ in the overall design process.

6.3 Randomization in Decentralized Neyman-Pearson Detection Systems

In this section we show that when the global false alarm probability is specified at some level α, then the global probability of detection β in (6.5) is maximized when the parameter γ is either 0 or 1. We begin by assuming that at some point (say poi: ⁺ P1 in Fig. 6.1) on the receiver operating characteristic (ROC) of an individual detector, the specified α is achieved when the randomization factor in (6.4) is 1. This means that the fusion rule at the fusion center is a k^* out of n rule. In general, a k out of n fusion rule is defined as follows : The global decision maker decides in favor of hypothesis H_1 when the number of local detectors that decide in favor of hypothesis H_1 is greater than or equal to k. It is shown in Appendix A that when the fusion center is a minimum average cost receiver, then under the identical detector assumption the optimum fusion rule is a k out of n rule. Let the local detection and false alarm probabilities at the point P1 be denoted by P_{D1} and P_{F1}, and let the local threshold that achieves these probabilities be denoted by τ_1 (see equation (4.5)). Now, assume that α is also achieved at two neighboring points P0 and P2 as shown in Fig. 6.1 with corresponding global detection probabilities β_0 and β_2 respectively. Let the local thresholds at the points P0 and P2 be denoted by τ_0 and τ_2. Now we prove the following theorem.

Theorem 6.1

In designing a globally optimum Neyman-Pearson decentralized parallel fusion system, the global probability of detection is maximized when the randomization factor at the fusion center with respect to the local decisions is either 0 or 1, i.e., system without randomization performs better than the ones employing randomization.
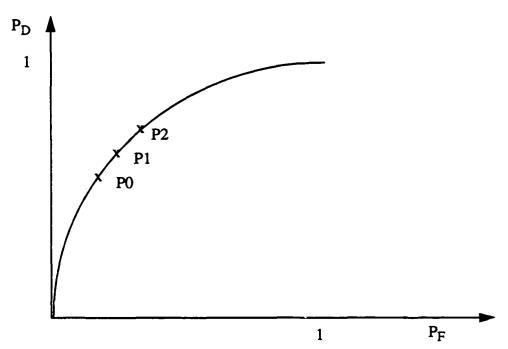
Fig. 6.1. Receiver operating characteristic of a local detector. The points P0, P1, and P2 are the points where the global false alarm probability is fixed at α .

Proof

To prove the theorem we need to show that $\beta_1 > \beta_0$ and $\beta_1 > \beta_2$. The proof will be carried out in two parts. In the first part we show that $\beta_1 > \beta_0$ and in the second we show that $\beta_1 > \beta_2$.

Part 1 : $\beta_1 > \beta_0$

Since at point P1, α is satisfied by a k out of n fusion rule with k = k*, the randomization factor in (6.1) is 1.Therefore (6.4) and (6.5) reduce to

$$\alpha = \sum_{j=k^*}^{n} \binom{n}{j} P_{F1}^j (1 - P_{F1})^{n-j} \tag{6.6}$$

$$\beta_1 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D1}^j (1 - P_{D1})^{n-j} \tag{6.7}$$

Now we show that for two positive numbers $0 < y_1 < 1$ and $0 < y_2 < 1$ such that $y_1 < y_2$, the following inequality is true

$$\sum_{j=k^*}^{n} \binom{n}{j} y_1^j (1 - y_1)^{n-j} < \sum_{j=k^*}^{n} \binom{n}{j} y_2^j (1 - y_2)^{n-j} \tag{6.8}$$

We prove the inequality in (6.8) by showing that the slope of the following function of y is positive

$$g(y) = \sum_{j=k^*}^{n} \binom{n}{j} y^j (1 - y)^{n-j} \tag{6.9}$$

Let us define the function $g_j(y)$ as

$$g_j(y) = y^j (1 - y)^{n-j} \tag{6.10}$$

132

The derivative of $g_j(y)$ is

$$\frac{d}{dy}g_j(y) = jy^{j-1}(1-y)^{n-j} - (n-j)y^j(1-y)^{n-j-1} \qquad (6.11)$$

Using (6.10) and (6.11), the derivative of the function $g(y)$ defined in (6.9) becomes

$$\frac{d}{dy}g(y) = S(y) = \sum_{j=k^*}^{n} \binom{n}{j}\{jy^{j-1}(1-y)^{n-j} - (n-j)y^j(1-y)^{n-j-1}\} \qquad (6.12)$$

It can be shown that the summation $S(y)$ in (6.12) simplifies to

$$S(y) = \frac{n!}{(k^*-1)!\,(n-k^*)!}y^{k^*-1}(1-y)^{n-k^*} \qquad (6.13)$$

The summation in (6.13) is positive for all the values of k^* when the values of y are in the range $0 < y < 1$. This result shows that the slope of the function $g(y)$ is positive indicating that $g(y)$ is an increasing function of y in the range $0 < y < 1$. Therefore, for $y_1 < y_2$, $g(y_1) < g(y_2)$ as claimed in (6.8).

From Fig. 6.1 we see that $P_{F0} < P_{F1}$. Therefore, using (6.8) we conclude that

$$\sum_{j=k^*}^{n} \binom{n}{j}P_{F0}^j(1-P_{F0})^{n-j} < \sum_{j=k^*}^{n} \binom{n}{j}P_{F1}^j(1-P_{F1})^{n-j} \qquad (6.14)$$

Using (6.6), (6.14) can be written as

$$\sum_{j=k^*}^{n} \binom{n}{j}P_{F0}^j(1-P_{F0})^{n-j} < \alpha \qquad (6.15)$$

In order to maintain a false alarm probability of α at point P0, a randomization term at $k = k^*-1$ has to be added to the left hand side of (6.15) so that equality is achieved. Therefore, the following decision rule is employed at the point P0

$$\phi(k) = \begin{cases} 1 & \text{when} \quad k > k^*-1 \\ \gamma_0 & \text{when} \quad k = k^*-1 \\ 0 & \text{otherwise} \end{cases} \tag{6.16}$$

Using (6.16) we evaluate the global false alarm and the global detection probabilities at the point P0 as

$$\alpha = \gamma_0 \binom{n}{k^*-1} P_{F0}^{k^*-1} (1-P_{F0})^{n-k^*+1} + \sum_{j=k^*}^{n} \binom{n}{j} P_{F0}^j (1-P_{F0})^{n-j} \tag{6.17}$$

$$\beta_0 = \gamma_0 \binom{n}{k^*-1} P_{D0}^{k^*-1} (1-P_{D0})^{n-k^*+1} + \sum_{j=k^*}^{n} \binom{n}{j} P_{D0}^j (1-P_{D0})^{n-j} \tag{6.18}$$

From (6.17), we find that γ_0 is given by

$$\gamma_0 = \frac{\alpha - \sum_{j=k^*}^{n} \binom{n}{j} P_{F0}^j (1-P_{F0})^{n-j}}{\binom{n}{k^*-1} P_{F0}^{k^*-1} (1-P_{F0})^{n-k^*+1}} \tag{6.19}$$

Substituting (6.19) into (6.18) we get

$$\beta_0 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D0}^j (1-P_{D0})^{n-j}$$

$$+ \quad \{\alpha - \sum_{j=k^*}^{n} \binom{n}{j} P_{F0}^j (1-P_{F0})^{n-j}\} \left(\frac{P_{D0}}{P_{F0}}\right)^{k^*-1} \left(\frac{1-P_{D0}}{1-P_{F0}}\right)^{n-k^*+1} \tag{6.20}$$

Substituting the value of α given in (6.6) into (6.20) we get

$$\beta_0 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D0}^j (1-P_{D0})^{n-j} \quad +$$

$$\{ \sum_{j=k^*}^{n} \binom{n}{j} [P_{F1}^j (1-P_{F1})^{n-j} - P_{F0}^j (1-P_{F0})^{n-j}] \} \left(\frac{P_{D0}}{P_{F0}}\right)^{k^*-1} \left(\frac{1-P_{D0}}{1-P_{F0}}\right)^{n-k^*+1} \tag{6.21}$$

134

Our objective now is to express the right hand side of (6.21) in terms of β_1. For this purpose we recall the following approximation based on Taylor series expansions of a function $g(y)$

$$g(y + \Delta y) \cong g(y) + g'(y)\,\Delta y \qquad (6.22)$$

$$g(y - \Delta y) \cong g(y) - g'(y)\,\Delta y \qquad (6.23)$$

In the rest of the chapter the approximation sign in (6.22) and (6.23) will be replaced by equality sign. Let us consider the function $g(y)$ given by (6.9). The derivative of this function is given by (6.12). Using (6.22) and (6.23) we get the following expansions of the function $g(y)$

$$g(y + \Delta y) = \sum_{j = k^*}^{n} \binom{n}{j} y^j (1-y)^{n-j} \quad +$$

$$\left\{ \sum_{j = k^*}^{n} \binom{n}{j} \left\{ j y^{j-1}(1-y)^{n-j} - (n-j) y^j (1-y)^{n-j-1} \right\} \right\} \Delta y \qquad (6.24)$$

$$g(y - \Delta y) = \sum_{j = k^*}^{n} \binom{n}{j} y^j (1-y)^{n-j} \quad -$$

$$\left\{ \sum_{j = k^*}^{n} \binom{n}{j} \left\{ j y^{j-1}(1-y)^{n-j} - (n-j) y^j (1-y)^{n-j-1} \right\} \right\} \Delta y \qquad (6.25)$$

Since $P_{F1} = P_{F0} + \Delta P_F$ (refer to Fig. 6.1), then by making use of (6.24) we obtain the fol-

lowing expansion

$$\sum_{j=k^*}^{n} \binom{n}{j} P_{F1}^j (1-P_{F1})^{n-j} = \sum_{j=k^*}^{n} \binom{n}{j} P_{F0}^j (1-P_{F0})^{n-j} +$$

$$\sum_{j=k^*}^{n} \binom{n}{j} (j P_{F0}^{j-1} (1-P_{F0})^{n-j} - (n-j) P_{F0}^j (1-P_{F0})^{n-j-1}) \Delta P_F$$

$$(6.26)$$

Substituting (6.26) into (6.21) we get

$$\beta_0 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D0}^j (1-P_{D0})^{n-j} + \left(\frac{P_{D0}}{P_{F0}}\right)^{k^*-1} \left(\frac{1-P_{D0}}{1-P_{F0}}\right)^{n-k^*+1}$$

$$x \quad \{ \sum_{j=k^*}^{n} \binom{n}{j} [j P_{F0}^{j-1} (1-P_{F0})^{n-j} - (n-j) P_{F0}^j (1-P_{F0})^{n-j-1}] \} \Delta P_F$$

$$(6.27)$$

Since $P_{D0} = P_{D1} - \Delta P_D$, then we can use (6.25) to obtain the following expansion

$$\sum_{j=k^*}^{n} \binom{n}{j} P_{D0}^j (1-P_{D0})^{n-j} = \sum_{j=k^*}^{n} \binom{n}{j} P_{D1}^j (1-P_{D1})^{n-j} -$$

$$\sum_{j=k^*}^{n} \binom{n}{j} (j P_{D1}^{j-1} (1-P_{D1})^{n-j} - (n-j) P_{D1}^j (1-P_{D1})^{n-j-1}) \Delta P_D$$

$$(6.28)$$

It can be easily recognized that the first term on the right hand side of (6.28) is the global probability of detection at the point P1 as given by (6.7). Substituting (6.28) into (6.27) and making use of (6.7) we get

$$\beta_0 = \beta_1 - S(P_{D1}) \Delta P_D + S(P_{F0}) \Delta P_F \left(\frac{P_{D0}}{P_{F0}}\right)^{k^*-1} \left(\frac{1-P_{D0}}{1-P_{F0}}\right)^{n-k^*+1}$$

$$(6.29)$$

136

where S(.) is given by (6.12). Making use of (6.13), equation (6.29) simplifies to

$$\beta_0 = \beta_1 - \frac{n!}{(k^* - 1)!\,(n - k^*)!} \quad \text{x}$$

$$\left[P_{D1}^{k^* - 1} (1 - P_{D1})^{n - k^*} \Delta P_D - \Delta P_F P_{F0}^{k^* - 1} (1 - P_{F0})^{n - k^*} \left(\frac{P_{D0}}{P_{F0}} \right)^{k^* - 1} \left(\frac{1 - P_{D0}}{1 - P_{F0}} \right)^{n - k^* + 1} \right]$$

$$(6.30)$$

In the limit when ΔP_D and ΔP_F approach zero, (6.30) becomes

$$\beta_0 = \beta_1 - \frac{n!}{(k^* - 1)!\,(n - k^*)!} P_{D1}^{k^* - 1} (1 - P_{D1})^{n - k^*} \left[\Delta P_D - \frac{1 - P_{D1}}{1 - P_{F1}} \Delta P_F \right]$$

$$(6.31)$$

But the following inequality is true for any operating point (P_F, P_D) on the receiver operating characteristic (see Appendix D for a proof)

$$\Delta P_D (1 - P_F) > \Delta P_F (1 - P_D) \tag{6.32}$$

Using this inequality, we see that the term between brackets on the right hand side of (6.31) is positive, and so we establish the following result

$$\beta_0 < \beta_1 \tag{6.33}$$

This concludes the first part of the proof. Now we consider the second part.

Part 2 : $\beta_1 > \beta_2$

From Fig. 6.1, we note that the false alarm probability P_{F2} at point P2 is greater than the false alarm probability P_{F1} at point P1. Therefore, using (6.8) we conclude that

$$\sum_{j = k^*}^{n} \binom{n}{j} P_{F2}^{j} (1 - P_{F2})^{n - j} > \sum_{j = k^*}^{n} \binom{n}{j} P_{F1}^{j} (1 - P_{F1})^{n - j} \tag{6.34}$$

137

Using (6.6) we can rewrite (6.39) as

$$\sum_{j=k^*}^{n} \binom{n}{j} P_{F2}^j (1 - P_{F2})^{n-j} > \alpha \qquad (6.35)$$

In order to maintain a false alarm probability of α at point P2, the outcome at $k = k^*$ on the left hand side of (6.35) is randomized so as to achieve equality in (6.35). Therefore, the following decision rule is employed at the point P2

$$\phi(k) = \begin{cases} 1 & \text{when} \quad k > k^* \\ \gamma_2 & \text{when} \quad k = k^* \\ 0 & \text{otherwise} \end{cases} \qquad (6.36)$$

Using (6.36) we evaluate the global false alarm and the global detection probabilities at the point P2 as

$$\alpha = \gamma_2 \binom{n}{k^*} P_{F2}^{k^*} (1 - P_{F2})^{n-k^*} + \sum_{j=k^*+1}^{n} \binom{n}{j} P_{F2}^j (1 - P_{F2})^{n-j} \qquad (6.37)$$

$$\beta_2 = \gamma_2 \binom{n}{k^*} P_{D2}^{k^*} (1 - P_{D2})^{n-k^*} + \sum_{j=k^*+1}^{n} \binom{n}{j} P_{D2}^j (1 - P_{D2})^{n-j} \qquad (6.38)$$

From (6.37), we find that γ_2 is given by

$$\gamma_2 = \frac{\alpha - \sum_{j=k^*+1}^{n} \binom{n}{j} P_{F2}^j (1 - P_{F2})^{n-j}}{\binom{n}{k^*} P_{F2}^{k^*} (1 - P_{F2})^{n-k^*}} \qquad (6.39)$$

Substituting (6.39) into (6.38) we get

$$\beta_2 = \sum_{j=k^*+1}^{n} \binom{n}{j} P_{D2}^j (1 - P_{D2})^{n-j}$$

$$+ \left\{ \alpha - \sum_{j=k^*+1}^{n} \binom{n}{j} P_{F2}^j (1 - P_{F2})^{n-j} \right\} \left(\frac{P_{D2}}{P_{F2}} \right)^{k^*} \left(\frac{1 - P_{D2}}{1 - P_{F2}} \right)^{n-k^*} \qquad (6.40)$$

138

Substituting the value of α given in (6.6) into (6.40) we get

$$\beta_2 = \sum_{j=k^*+1}^{n} \binom{n}{j} P_{D2}^{j} (1-P_{D2})^{n-j} + \left(\frac{P_{D2}}{P_{F2}}\right)^{k^*} \left(\frac{1-P_{D2}}{1-P_{F2}}\right)^{n-k^*}$$

$$\times \quad \left\{ \sum_{j=k^*}^{n} \binom{n}{j} P_{F1}^{j} (1-P_{F1}^{j})^{n-j} - \sum_{j=k^*+1}^{n} \binom{n}{j} P_{F2}^{j} (1-P_{F2}^{j})^{n-j} \right\} \qquad (6.41)$$

As we did in the first part, here our objective is to express β_2 in terms of β_1. Since $P_{F1} = P_{F2} - \Delta P_F$, then by making use of (6.25) we obtain the following expansion

$$\sum_{j=k^*}^{n} \binom{n}{j} P_{F1}^{j} (1-P_{F1})^{n-j} = \sum_{j=k^*}^{n} \binom{n}{j} P_{F2}^{j} (1-P_{F2})^{n-j}$$

$$\sum_{j=k^*}^{n} \binom{n}{j} \left(j P_{F2}^{j-1} (1-P_{F2})^{n-j} - (n-j) P_{F2}^{j} (1-P_{F2})^{n-j-1} \right) \Delta P_F$$

$$(6.42)$$

Substititing (6.42) into (6.41) we get

$$\beta_2 = \sum_{j=k^*+1}^{n} \binom{n}{j} P_{D2}^{j} (1-P_{D2})^{n-j} + \left(\frac{P_{D2}}{P_{F2}}\right)^{k^*} \left(\frac{1-P_{D2}}{1-P_{F2}}\right)^{n-k^*} \quad \times$$

$$\left\{ \binom{n}{k} P_{F2}^{k^*} (1-P_{F2})^{n-k^*} - S(P_{F2}) \Delta P_F \right\} \qquad (6.43)$$

where S(.) is given by (6.12). In obtaining (6.43) we have made use of the fact that

$$\sum_{j=k^*}^{n} \binom{n}{j} y^{j} (1-y)^{n-j} - \sum_{j=k^*+1}^{n} \binom{n}{j} y^{j} (1-y)^{n-j} = \binom{n}{k^*} y^{j} (1-y)^{n-k^*}$$

$$(6.44)$$

Equation (6.43) can also be written as

139

$$\beta_2 = \sum_{j=k^*+1}^{n} \binom{n}{j} P_{D2}^{j} (1 - P_{D2})^{n-j} + \binom{n}{k^*} P_{D2}^{k^*} (1 - P_{D2})^{n-k^*}$$

$$-S(P_{F2}) \left(\frac{P_{D2}}{P_{F2}}\right)^{k^*} \left(\frac{1-P_{D2}}{1-P_{F2}}\right)^{n-k^*} \Delta P_F \qquad (6.45)$$

Using the expression for S(.) given by (6.13), equation (6.45) reduces to

$$\beta_2 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D2}^{j} (1 - P_{D2})^{n-j}$$

$$-\frac{n!}{(k^*-1)!(n-k^*)!} P_{F2}^{k^*-1} (1-P_{F2})^{n-k^*} \left(\frac{P_{D2}}{P_{F2}}\right)^{k^*} \left(\frac{1-P_{D2}}{1-P_{F2}}\right)^{n-k^*} \Delta P_F$$

$$(6.46)$$

Now expand the first term on the right hand side of (6.46) about the point P_{D1}

$$\beta_2 = \sum_{j=k^*}^{n} \binom{n}{j} P_{D1}^{j} (1 - P_{D1})^{n-j} - \frac{n!}{(k^*-1)!(n-k^*)!} \frac{P_{D2}^{k^*}}{P_{F2}} (1-P_{D2})^{n-k^*} \Delta P_F$$

$$+ \sum_{j=k^*}^{n} \binom{n}{j} [j P_{D1}^{j-1} (1 - P_{D1})^{n-j} - (n-j) P_{D1}^{j} (1 - P_{D1})^{n-j-1}] \Delta P_D$$

$$(6.47)$$

The first term on the right hand side of (6.47) is β_1. Using (6.13), equation (6.47) becomes

$$\beta_2 = \beta_1 - \frac{n!}{(k^*-1)!(n-k^*)!} \frac{P_{D2}^{k^*}}{P_{F2}} (1-P_{D2})^{n-k^*} \Delta P_F$$

$$+ \frac{n!}{(k^*-1)!(n-k^*)!} P_{D1}^{k^*-1} (1-P_{D1})^{n-k^*} \Delta P_D \qquad (6.48)$$

As point P2 becomes closer and closer to point P1, equation (6.48) can be expressed as

$$\beta_2 = \beta_1 - \frac{n!}{(k^* - 1)!(n - k^*)!} P_{D1}^{k^* - 1} (1 - P_{D1})^{n - k^*} \{\frac{P_{D1}}{P_{F1}}\Delta P_F - \Delta P_D\}$$

(6.49)

But for any point (P_F, P_D) on the receiver operating characteristic of an individual sensor, the following inequality is true (see Fig. 6.2 and Appendix D)

$$P_D \Delta P_F > P_F \Delta P_D$$

(6.50)

Making use of this inequality, we see that the second term on the right hand side of (6.49) is positive which indicates that

$$\beta_2 < \beta_1$$

(6.51)

This completes the proof that randomization of the observations at the fusion center is not necessary when the objective is to design a globally optimum Neyman-Pearson decentralized detection system.

6.4 Summary

In this chapter we addressed the problem of randomization with respect to the local decisions at the fusion center when the objective is to design an optimum Neyman-Pearson decentralized detection system. We found that for a prespecified value of the global probability of false alarm the global probability of detection is maximized when there is no randomization, i.e., the randomization factor is either 0 or 1. This means that a k out of n fusion rule is the fusion rule that maximizes the global probability of detection ,i.e., the global probability of detection has (k+1) local maximum points for the specified false alarm probability. One remaining aspect of the problem is to determine on theoretical basis the particular fusion rule with the largest probability of detection.
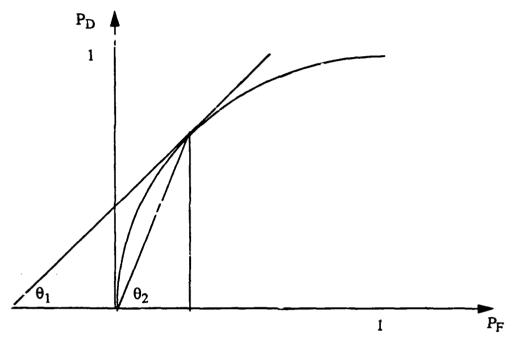
Fig. 6.2. Due to the concavity of the receiver operating characteristic, $\theta_1 < \theta_2$ where $\theta_1 = dP_D / dP_F$ is the slope and $\theta_2 = P_D / P_F$.

CHAPTER 7

SUMMARY AND SUGGESTIONS FOR FUTURE WORK

7.1 Summary

In this dissertation, we considered the design and performance evaluation of distributed detection networks mainly from a Bayesian viewpoint. In a distributed detection system, a group of local detectors process the observations they receive about the status of a certain phenomenon, and transmit their decisions to a fusion center where the final decision is made. Due to constraints on the channel capacities linking the local detectors to the global decision maker, the local detectors compress their local observations to a hard or a soft decision and transmit this decision to the fusion center where the global decision is made. The design issues related to the distributed detection systems involve specifying the local decision rules and the global decision rule. We presented a computationally efficient approach to the design of decentralized Bayesian detection systems. This procedure is based upon an alternate representation of the minimum average cost in terms of a modified form of the Kolmogorov variational distance. We demonstrated the utility of our approach by applying it to the design and performance evaluation of four decentralized detection structures. In all these structures, the design of the optimum systems reduced to the optimization of a single function of a certain number of variables. This optimization is performed once and does not need an exhaustive search. Two methods that deal with the design of binary decentralized Bayesian detection systems are reported in the literature. In the first method, the fusion center is fixed and a set of n coupled nonlinear equations are solved to determine the n local thresholds. This has to be repeated for all the permissible fusion rules. The solution with the smallest overall cost is finally selected as the optimum system. The exponential growth of the number of fusion rules to be searched makes the use of this method impractical. The other method is the person-by-person optimization

procedure which need not yield the globally optimum solution. This procedure requires a simultaneous solution of (2^n+n) coupled nonlinear equations for the binary hypothesis testing problem. Our design procedure requires the minimization of a single function of n variables where n is the number of local detectors. This is computationally simpler and efficient optimization algorithms can be employed to design relatively large decentralized detection structures.

The performance degradation of distributed detection systems relative to the centralized system was determined. It was also shown that the performance of the decentralized system approaches the performance of the centralized system very quickly as a function of the number of quantization levels. In addition, the performance of the optimum systems was compared to the performance of suboptimum systems designed based on criteria other than the global optimum Beyesian cost.

We derived upper and lower bounds on the minimum probability of error and the minimum average cost for optimum receivers. Some of the bounds on the minimum average cost are generalizations of the known bounds on the minimum probability of error in terms of the Bhattacharyya bound and the Chernoff bound. Furthermore, we derived a new upper bound on the probability of error which is tighter than the previously available bounds. This bound was applied to design a nearly optimum decentralized detection system. Tight Lower bounds on the probability of error and the minimum average cost were also derived in terms of the new upper bound.

We addressed the problem of randomization with respect to the local decisions at the fusion center when the objective is to design an optimum Neyman-Pearson decentralized detection system. We found that for a prespecified value of the global probability of false alarm, the global probability of detection is maximized when there is no randomization, i.e., the randomization factor is either 0 or 1.

7.2 Suggestions for Future Work

In this section, we discuss some of the problems related to the topics treated in this dissertation and may be pursued in the future.

1) In Chapter 4, we employed the new upper bound derived in Chapter 3 to design a nearly optimum decentralized detection system. A generalization of the problem is to employ the new upper bound for the design of an M-level quantizer for hypothesis testing.

2) In decentralized detection systems, usually there are capacity constraints on the channels linking the local detectors to the destination. An interesting problem would be to examine from an information theoretic viewpoint, the design of decentralized detection systems when capacity constraints are placed on the channels. For example, the output entropy may be restricted to be less than some prespecified value. One then needs to find the fusion rule and the local decision rules under this constraint.

3) In Chapter 6, we investigated the role of the randomization factor in the decentralized Neyman-Pearson detection problem. We considered the case when the local detectors were identical and they employed identical decision rules. We found that randomization was not necessary in the design of the optimum system. Two problems related to this result arise. The first is to determine, based on a theoretical basis, the fusion rule with the highest probability of detection for the specified probability of false alarm. The second is to show that this result is true for nonidentical local detectors.

REFERENCES

[1] H.L.Van Trees, *Detection, Estimation and Modulation Theory*, Vol. 1, Wiley, 1968.

[2] H.V.Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, N.Y. 1988.

[3] R.E.Ziemer and W.H.Tranter, *Principles of Communications, Systems, Modulation, and Noise*, Houghton Mifflin, 1985.

[4] J.N.Tsitsiklis and M.Athans, "On the Complexity of Decentralized Decision Making and Detection Problems," *IEEE Transactions on Automatic Control*, Vol. 30, pp. 440-446, May 1985.

[5] R.R.Tenney and N.R.Sandell, " Structures of Distributed Decisionmaking," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 11, pp. 517-526, Aug. 1981.

[6] R.R.Tenney and N.R.Sandell, " Strategies for Distributed Decisionmaking," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 11, pp. 527-538, Aug. 1981.

[7] R.R.Tenney and N.R.Sandell, " Detection with Distributed Sensors," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 17, pp. 501-510, July 1981.

[8] F.A.Sadjadi, " Hypothesis Testing in a Distributed Environment," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 22, pp. 134-137, March 1987.

[9] Z.Chair and P.K.Varshney, "Optimal Data Fusion in Multiple Detection Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 22, pp. 98-101, Jan. 1986.

[10] I.Hoballah and P.K.Varshney, "Decentralized Bayesian Signal Detection," *IEEE Transactions on Information Theory*, Vol. 35, pp. 995-1000, Sep. 1989.

[11] Z.B.Tang,K.R.Pattipati,and D.L.Kleinman, "An Algorithm for Determining the Decision Thresholds in a Distributed Detection Problem," Preprint, University of Connecticut.

[12] A.R.Reibman and L.W.Nolte, "Optimal Detection and Performance of Distributed Sensor Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 23, pp.

24-30, Jan. 1987.

[13] S.C.A.Thomopoulos,R.Viswanathan,and D.K.Bougoulias, "Optimal Distributed Decision Fusion," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 25, pp. 761-765, Sep. 1989.

[14] Z.Chair and P.K.Varshney, "Distributed Bayesian Hypothesis Testing with Distributed Data Fusion," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 18, pp. 695-699, Sep. /Oct. 1988.

[15] R.Srinivasan, "Distributed Radar Detection Theory," *IEE Proceedings*, Vol. 133-F, pp. 55-60, Feb. 1986.

[16] I.Y.Hoballah, " On the Design and Optimization of Distributed Signal Detection and Parameter Estimation Systems," PhD Dissertation, Syracuse University, 1986.

[17] S.C.A.Thomopoulos,R.Viswanathan,and D.K.Bougoulias, " Optimal Decision Fusion in Multiple Sensor Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 23, pp. 644-654, Sep. 1987.

[18] J.N.Tsitsiklis, " Extremal Properties of Likelihood Ratio Quantizers," Preprint, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Nov. 1989.

[19] A.Reibman, " Performance and Fault-Tolerance of Distributed Detection Networks," PhD Dissertation, Department of Electrical Engineering, Duke University, Durham, North Carolina, 1987.

[20] P.Willet and D.Warren, " The Suboptimality of Randomized Tests in Distributed and Quantized Detection Systems," Preprint, Nov. 1989.

[21] R.Viswanathan,S.C.A.Thomopoulos,and R.Tumuluri, "Optimal Serial Distributed Decision Fusion," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 24, pp. 366-376, July 1988.

[22] L.K.Ekchian and R.R.Tenney, "Detection Networks," Proceedings of the 21st IEEE Conference on Decision and Control, Orlando, Fla., pp. 686-691, Dec. 1982.

[23] A.R.Reibman and L.W.Nolte, "Design and Performance Comparison of Distributed Detection Networks," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 23, pp. 789-797, Nov. 1987.

[24] J.N.Tsitsiklis, "Decentralized Detection," Technical Report LIDS-P-1913, M.I.T., Sep. 1989.

[25] P.K.Varshney, " Decentralized Bayesian Detection Theory," Invited Chapter to appear in the Book *Large - Scale Stochastic Systems: Detection, Estimation, Stability and Control*, S.G.Tzafestas and K.Watanabe, Editors, Marcel Dekker Inc.

[26] J.D.Papastavrou and M.Athans, "On Optimal Distributed Decision Architectures in a Hypothesis Testing Environment," Technical Report LIDS-P-1928, M.l.T., Jan. 1990.

[27] P.Willet and D.Warren, "Randomization and Scheduling in Decentralized Detection," Proceedings of the 27 [th] Annual Conference on Communication, Control, and Computing, Sep. 1989.

[28] E.Drakopoulos and C.C.Lee, "Optimum Multisensor Fusion of Correlated Local Decisions," Preprint, Department of Electrical Engineering and Computer Science, Northwestern University.

[29] G.Polychronopoulos and J.Tsitsiklis, " Explicit Solutions for Some Simple Decentralized Detection Problems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 26, pp. 282-292, March 1990.

[30] S.M.Ali and S.D.Silvey, " A General Class of Coefficients of Divergence of One Distribution From Another," *J.Royal Stat. Soc., Series B*, Vol. 28, pp. 131-143, 1966.

[31] T.Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions on Communication Technology*, Vol. 15, pp. 52-60, Feb. 1967.

[32] H.V.Poor and J.B.Thomas, "Applications of Ali-Silvey Distance Measures in the Design of Generalized Quantizers for Binary Decision Systems," *IEEE Transactions on Communications*, Vol. 25, pp. 893-900, Sep. 1977.

[33] H.V.Poor, " Fine Quantization in Signal Detection and Estimation," *IEEE Transac-*

tions on Information Theory, Vol. 34, pp. 960-972, Sep. 1988.

[34] G.R.Benitz and J.A.Bucklew, " Asymptotically Optimal Quantizers for Detection of I.I.D. Data," *IEEE Transactions on Information Theory*, Vo. 35, pp. 316-325, March 1989.

[35] M.Longo,T.Lookabaugh,and R.M.Gray, "Quantization for Decentralized Hypothesis Testing Under Communication Constraints," *IEEE Transactions on Informat.. ∴ Theory*, Vol.36, pp. 241-255, March 1990.

[36] C.C.Lee and J.J.Chao, "Optimam Local Decision Space Partitioning for Distributed Detection," *IEEE Transactions on Aerospace cnd Electronic Systems*, Vol. 25, pp. 536-544, July 1989.

[37] A.T.Toussaint, " Comments on the Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions on Communications*, Vol. 20, pp. 485, June 1972.

[38] T.L.Grettenberg, " Signal Selection in Communication and Radar Systems," *IEEE Transactions on Information Theory*, Vol. 9, pp. 265-272, Oct. 1963.

[39] A.Gersho, " Principles of Quantization," *IEEE Transactions on Circuits and Systems*, Vol. 25, pp. 427-436, July 1978.

[40] P.Noll and R.Zelinski, " Bounds and Quantizer Performance in the Low Bit Rate Region," *IEEE Transactions on Communications*, Vol. 26, pp. 300-304, Feb. 1978.

[41] N.Farvardin and J.W.Modestino, " Optimum Quantizer Performance for a class of NonGaussian Memoryless Sources," *IEEE Transactions on Information Theory*, Vol. 30, pp. 485-497, May 1984.

[42] S.A.Kassam, " Optimum Quantization for Signal Detection," *IEEE Transactions on Communications*, Vol. 25, pp. 479-484, May 1977.

[43] B.Aazhang and H.V.Poor, " On Optimum and Nearly Optimum Data Quantization for Signal Detection," *IEEE Transactions on Communications*, Vol. 32, pp. 745-751, Jul. 1984.

[44] D.Middleton, *An Introduction to Statistical Communication Theory*, McGraw Hill, 1960.

[45] T.L.Gabriele, " Information Criteria for Threshold Determination, " *IEEE Transactions on Information Theory*, Vol. 12 ,pp. 484-486, Oct. 1966.

[46] A.B.Martinez, " A Comparison of Bayesian and Information-Optimal Detection Thresholds," Proceedings of the 1987 American Control Conference, Minneapolis, MN, pp. 1951-1956, June 1987.

[47] I.Y.Hoballah and P.K.Varshney, " An Information Theoretic Approach to the Distributed Detection Problem," *IEEE Transactions on Information Theory*, Vol. 35, pp. 988-994, Sep. 1989.

[48] D.Kazakos,V.Vannicola,and M.Wicks, "On Optimal Quantization for Detection," Proceedings of the 1989 Conference on Information Science and Systems, March 1989.

[49] J.Max, " Quantizing for Minimum Distortion," *IRE Transactions on Information Theory*, Vol. 6, pp. 7-12, March 1960.

[50] B.Picinbono and P.Duvaut, " Optimum Quantization for Detection," *IEEE Transactions on Communications*, Vol. 36, pp. 1254-1258, Nov. 1988.

[51] D.E.Boekee and J.C.A.Van der Lubbe, "Some Aspects of Error Bounds in Feature Selection," *Pattern Recognition*, Vol. 11, pp. 353-360, 1979.

[52] H.Chernoff, "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on a Sum of Observations," *Ann.Math. Stat.*, Vol. 23, pp. 493-507, 1952.

[53] M.E.Hellman, "Probability of Error, Equivocation, and Chernoff Bound," *IEEE Transactions on Information Theory*, Vol. 16, pp. 368-372, July 1970.

[54] T.Ito, "Approximate Error Bounds in Pattern Recognition," *Machine Intelligence*, Vol. 7, pp. 369-376, 1972.

[55] C.H.Chen, "Theoretical Comparison of a Class of Feature Selection Criteria in Pattern Recognition," *IEEE Transactions on Computers*, Vol. C-20, pp. 1054-1056, Sep. 1971.

[56] P.A.Devijver, "On a New Class of Bounds on Bayes Risk in Multihypothesis Pattern Recognition," *IEEE Transactions on Computers*, Vol. C-23, pp. 70-80, Jan. 1974.

[57] D.Kazakos, " New Error Bounds and Optimum Quantization for Multisensor Distributed Signal Detection," Preprint, Electrical Engineering Department, University of Virginia.

[58] J.Tsitsiklis, " Decentralized Detection by a Large Number of Sensors," *Mathematics of Control, Signals, and Systems*, 1, pp. 167-182, 1988.

[59] C.H.Chen, "On Information and Distance Meaures, Error Bounds, and Feature Selection,"*Information Science*, Vol. 10, pp. 159-171, 1976.

[60] D.E.Boekee and J.C.Ruitenbeek, "A Class Of Lower Bounds on The Bayesian Probability of Error," *Information Science*, pp. 32-35, 1981.

[61] G.T.Toussaint, "On the Divergence Between Two Distributions and the Probability of Misclassification of Several Decision Rules," Proceedings of the Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, Aug. 1974.

[62] R.O.Duda and P.E.Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, Inc., 1973.

[63] P.R.Krishnaiah and L.N.Kanal, *Handbook of Statistics 2, Classification Pattern Recognition and Reduction of Dimensionality*, North-Holland, 1982.

[64] R.Alshwede and I.Csiszar, " Hypothesis Testing with Communication Constraints," *IEEE Transactions on Information Theory*, Vol. 32, pp. 533-542, Jul. 1986.

[65] T.S.Han, " Hypothesis Testing with Multiterminal Data Compression," *IEEE Transactions on Information Theory*, Vol. 33, pp. 759-772, Nov. 1987.

[66] T.J.Flynn and R.M.Gray, " Encoding of Correlated Observations," *IEEE Transactions on Information Theory*, Vol. 33, pp. 773-787, Nov. 1987.

[67] R.Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, 1987.

[68] C.C.Lee and J.B.Thomas, " Detectors for Multinomial Input," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 19, pp. 288-296, March 1983.

[69] G.R.Walsh, *Methods of Optimization*, John Wiley and Sons, 1975.

Appendix A

In this appendix, we derive the algorithms based on which the fusion center combines the decisions received from the various local detectors in order to make the global decision. We consider both the hard decision and the soft decision cases, i.e., the cases with $M = 2$ and $M > 2$.

First Case : M = 2

Let P_{Di} and P_{Fi} be the probabilities of detection and false alarm at the local detector LD_i. The local decision z_i has the following conditional distributions

$$f_0(z_i) = P_{Fi}^{z_i}(1-P_{Fi})^{1-z_i} \tag{A.1}$$

$$f_1(z_i) = P_{Di}^{z_i}(1-P_{Di})^{1-z_i} \tag{A.2}$$

The global decision u_0 is made at the fusion center based on the decision vector $\underline{U} = [z_1 \, z_2 \, ... \, z_n]$ by performing the likelihood ratio test

$$\Lambda(\underline{U}) = \frac{P(\underline{U}|H_1)}{P(\underline{U}|H_0)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta \tag{A.3}$$

Taking the logarithm of both sides of (A.3) we get

$$\log\Lambda(\underline{U}) \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log\eta \tag{A.4}$$

Using (A.1), (A.2), and making use of the independence assumption between the local decisions z_i, we get

$$\Lambda(\underline{U}) = \frac{\displaystyle\prod_{i=1}^{n} P_{Di}^{z_i}(1-P_{Di})^{1-z_i}}{\displaystyle\prod_{i=1}^{n} P_{Fi}^{z_i}(1-P_{Fi})^{1-z_i}} \qquad (A.5)$$

Equation (A.5) can also be written as

$$\Lambda(\underline{U}) = \prod_{i=1}^{n} \left(\frac{P_{Di}}{P_{Fi}}\right)^{z_i}\left(\frac{1-P_{Di}}{1-P_{Fi}}\right)^{1-z_i} \qquad (A.6)$$

Taking the logarithm of both sides of (A.6) we get

$$\log\Lambda(\underline{U}) = \sum_{i=1}^{n} \log\left[\left(\frac{P_{Di}}{P_{Fi}}\right)^{z_i}\left(\frac{1-P_{Di}}{1-P_{Fi}}\right)^{1-z_i}\right] \qquad (A.7)$$

Expanding the logarithm of the product term on the right hand side of (A.7) and simplifying we get

$$\log\Lambda(\underline{U}) = \sum_{i=1}^{n} \left[z_i\log\left(\frac{P_{Di}}{P_{Fi}}\right) + (1-z_i)\log\left(\frac{1-P_{Di}}{1-P_{Fi}}\right)\right] \qquad (A.8)$$

Combining terms in (A.8) we obtain

$$\log\Lambda(\underline{U}) = \sum_{i=1}^{n} \log\left(\frac{1-P_{Di}}{1-P_{Fi}}\right) + \sum_{i=1}^{n} z_i\log\frac{P_{Di}(1-P_{Fi})}{P_{Fi}(1-P_{Di})} \qquad (A.9)$$

Substituting (A.9) into (A.4) and rearranging terms we get

$$\sum_{i=1}^{n} z_i\log\frac{P_{Di}(1-P_{Fi})}{P_{Fi}(1-P_{Di})} \underset{\substack{< \\ H_0}}{\overset{\substack{H_1 \\ >}}{}} \log\eta + \sum_{i=1}^{n} \log\left(\frac{1-P_{Fi}}{1-P_{Di}}\right) \qquad (A.10)$$

The equivalent test in (A.10) suggests that each received local decision z_i ; $z_i = 0, 1$ is weighted by a factor that depends upon the probabilities of detection and false alarm cor-

responding to that detector. The sum is then compared to a threshold.

Under the identical sensor assumption, i.e., when $P_{Di} = P_D$ and $P_{Fi} = P_F$; $i = 1, 2,...,$ n, the left hand side and the right hand side of (A.10) become

$$LHS = \log \frac{P_D(1 - P_F)}{P_F(1 - P_D)} \sum_{i=1}^{n} z_i \qquad (A.11)$$

$$RHS = \log \eta + n \log \left(\frac{1 - P_F}{1 - P_D} \right) = \log \eta \left(\frac{1 - P_F}{1 - P_D} \right)^n \qquad (A.12)$$

Therefore, equation (A.10) can be expressed as

$$\sum_{i=1}^{n} z_i \quad \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \quad \frac{\log \eta \left(\frac{1 - P_F}{1 - P_D} \right)^n}{\log \frac{P_D(1 - P_F)}{P_F(1 - P_D)}} \qquad (A.13)$$

Let K be the random variable that represents the number of sensors that decide in favor of H_1. It is clear that K is nothing but the sum on the left hand side of (A.13). This sum takes on the values 0, 1, ..., n. A k out of n fusion rule is defined as follows : Decide H_1 if the number of sensors that decide in favor of H_1 is at least k, i.e., decide H_1 if

$$\sum_{i=1}^{n} z_i \geq k, k = 1, 2, ..., n \qquad (A.14)$$

It is clear from (A.13) and (A.14) that the fusion rule at the fusion center is a k out of n fusion rule. The case with k = 0 has been excluded because it corresponds to the trivial case of always deciding H_1. Moreover, the case k = 0 violates the validity of (A.13) since the right hand side of (A.13) is a positive quantity and cannot be 0 or negative.

Second Case : M > 2

For the case when the number of quantization levels $M > 2$, we restrict our study to the case of identical detectors. Here we make use of the multinomial distributions derived in Chapter 4. Using (4.37) and (4.38), we see that

$$\Lambda(\underline{U}) = \frac{P_{11}^{x_1} P_{21}^{x_2} \cdots P_{M1}^{x_M}}{P_{10}^{x_1} P_{20}^{x_2} \cdots P_{M0}^{x_M}} \tag{A.15}$$

Equation (A.15) can also be written as

$$\Lambda(\underline{U}) = \left(\frac{P_{11}}{P_{10}}\right)^{x_1} \left(\frac{P_{21}}{P_{20}}\right)^{x_2} \cdots \left(\frac{P_{M1}}{P_{M0}}\right)^{x_M} \tag{A.16}$$

Taking the logarithm of both sides of (A.16) we get

$$\log \Lambda(\underline{U}) = x_1 \log\left(\frac{P_{11}}{P_{10}}\right) + x_2 \log\left(\frac{P_{21}}{P_{20}}\right) + \ldots + x_M \log\left(\frac{P_{M1}}{P_{M0}}\right) \tag{A.17}$$

Substituting (A.17) into (A.4) we obtain the test

$$x_1 \log\left(\frac{P_{11}}{P_{10}}\right) + x_2 \log\left(\frac{P_{21}}{P_{20}}\right) + \ldots + x_M \log\left(\frac{P_{M1}}{P_{M0}}\right) \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \log \eta \tag{A.18}$$

The equivalent test in (A.18) indicates that the fusion center makes the final decision by counting the number of sensors x_k that decide in favor of symbol a_k ; $k = 1, 2, \ldots, M$, weight x_k by a factor that depends upon the probabilities of occurrence of symbol a_k under hypotheses H_0 and H_1, sum over all values of x_k and compare the sum to a threshold.

Appendix B

In this appendix we provide a derivation of equation (4.18). The possible fusion rules for a three-sensor system are the AND fusion rule, the OR fusion rule and the MA-JORITY logic fusion rule. The cost corresponding to each one of these fusion rules is given by (4.14), (4.15) and (4.16) respectively. Since the fusion center is a minimum average cost receiver, it implements the 'best' one of the fusion rules, i.e., the fusion rule with the smallest cost. The cost corresponding to this best fusion rule is given by

$$R_3 = min \; (min \; (R_{AND}, R_{OR}), R_{MAJ}) \qquad (B.1)$$

The minimum of two quantities a and b is given by

$$min \, (a, b) \; = \; \frac{1}{2} \, (a + b) - \frac{1}{2} |a - b| \qquad (B.2)$$

Making use of (B.2), (B.1) becomes

$$R_3 = \frac{1}{2} \, (min \, (R_{AND}, R_{OR}) + R_{MAJ}) - \frac{1}{2} |min \, (R_{AND}, R_{OR}) - R_{MAJ}| \qquad (B.3)$$

Again applying (B.2) to (B.3) we obtain

$$R_3 = \frac{1}{2} \left[\frac{1}{2} \, (R_{AND} + R_{OR}) - \frac{1}{2} |R_{AND} - R_{OR}| + R_{MAJ} \right]$$

$$- \frac{1}{2} \left| \frac{1}{2} \, (R_{AND} + R_{OR}) - \frac{1}{2} |R_{AND} - R_{OR}| - R_{MAJ} \right| \qquad (B.4)$$

Simplifying (B.4) we get

$$R_3 = \frac{1}{4}\left[R_{AND} + R_{OR} + 2R_{MAJ}\right] - \frac{1}{4}\left|R_{AND} - R_{OR}\right|$$

$$-\frac{1}{4}\left|R_{AND} + R_{OR} - 2R_{MAJ} - \left|R_{AND} - R_{OR}\right|\right| \tag{B.5}$$

For simplicity of presentation we use the following notation

$$Q_1 = R_{AND} + R_{OR} + 2R_{MAJ} \tag{B.6}$$

$$Q_2 = R_{AND} - R_{OR} \tag{B.7}$$

$$Q_3 = R_{AND} + R_{OR} - 2R_{MAJ} \tag{B.8}$$

Substituting (4.14), (4.15), and (4.16) into (B.6) we get

$$Q_1 = C_{00}\pi_0\left[4 + 2P_F^3 - 3P_F - 3P_F^2\right] + C_{10}\pi_0\left[-2P_F^3 + 3P_F + 3P_F^2\right]$$

$$+ \quad C_{01}\pi_1\left[3P_M + 3P_M^2 - 2P_M^3\right] + C_{11}\pi_1\left[4 + 2P_M^3 - 3P_M - 3P_M^2\right]$$

$$\tag{B.9}$$

Adding and subtracting P_F to the first two terms on the right hand side of (B.9) and P_M to the last two terms we get

$$Q_1 = C_{00}\pi_0 [4 - 4P_F + 2P_F^3 - 3P_F^2 + P_F] + C_{10}\pi_0 [4P_F - P_F - 2P_F^3 + 3P_F^2]$$

$$+ \quad C_{01}\pi_1 [4P_M - P_M - 2P_M^3 + 3P_M^2] + C_{11}\pi_1 [4 - 4P_M + 2P_M^3 - 3P_M^2 + P_M] \tag{B.10}$$

Dividing (B.10) by 4 and expanding terms we get

$$\frac{Q_1}{4} = C_{00}\pi_0 (1 - P_F) + C_{10}\pi_0 P_F + C_{01}\pi_1 P_M + C_{11}\pi_1 (1 - P_M)$$

$$+ \quad \frac{C_{00}\pi_0}{4} [2P_F^3 - 3P_F^2 + P_F] - \frac{C_{10}\pi_0}{4} [P_F - 3P_F^2 + 2P_F^3]$$

$$- \frac{C_{01}\pi_1}{4} [P_M - 3P_M^2 + 2P_M^3] + \frac{C_{11}\pi_1}{4} [2P_M^3 - 3P_M^2 + P_M] \tag{B.11}$$

It is to be observed that the first term on the right hand side of (B.11) is the cost corresponding to a single sensor system. This cost was defined in (4.12). Rearranging terms in (B.11) we get

$$Q_1 = R_1 - \frac{\pi_0}{4} P_F (1 - 3P_F + 2P_F^2) (C_{10} - C_{00})$$

$$- \frac{\pi_1}{4} P_M (1 - 3P_M + 2P_M^2) (C_{01} - C_{11}) \tag{B.12}$$

Now we consider (B.7). Substituting (4.14), (4.15), and (4.16) into (B.7) we get

$$Q_2 = C_{00}\pi_0 [3P_F - 3P_F^2] + C_{10}\pi_0 [3P_F^2 - 3P_F]$$

$$+ \quad C_{01}\pi_1 [3P_M - 3P_M^2] + C_{11}\pi_1 [3P_M^2 - 3P_M] \tag{B.13}$$

Rearranging terms in (B.13) we get

$$Q_2 = 3\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) - 3\pi_0 P_F (1 - P_F) (C_{10} - C_{00})$$

$$\tag{B.14}$$

Dividing both sides of (B.14) and taking the absolute value we obtain

$$\frac{|Q_2|}{4} = \frac{3}{4} |\pi_1 P_M (1 - P_M) (C_{01} - C_{11}) - \pi_0 P_F (1 - P_F) (C_{10} - C_{00})|$$

$$\tag{B.15}$$

Substituting (4.14), (4.15), and (4.16) into (B.8) we get

$$Q_3 = C_{00}\pi_0 [-6P_F^3 + 9P_F^2 - 3P_F] + C_{10}\pi_0 [6P_F^3 + 3P_F - 9P_F^2]$$

$$+ \quad C_{01}\pi_1 [6P_M^3 + 3P_M - 9P_M^2] + C_{11}\pi_1 [-6P_M^3 + 9P_M^2 - 3P_M] \tag{B.16}$$

Rearranging terms in (B.16) and dividing by 4 we get

$$\frac{Q_3}{4} = \frac{3\pi_0}{4} P_F (1 - 3P_F + 2P_F^2) (C_{10} - C_{00}) + \frac{3\pi_1}{4} P_M (1 - 3P_M + 2P_M^2) (C_{01} - C_{11})$$

$$\tag{B.17}$$

Substituting (B.12), (B.15), and (B.17) into (B.5) we readily obtain (4.18).

Appendix C

In this appendix we derive the relationships between the regional threshold η_i and the local thresholds τ_{2i-1} and τ_{2i} for the regional decision maker RD_i in the hierarchical system S_4. This relationship is important in carrying out the design of the optimum system. We start with (5.32) which is repeated here for convenience

$$L_i = A_{2i-1}z_{2i-1} + A_{2i}z_{2i} \tag{C.1}$$

Recall that the random variables z_{2i-1} and z_{2i} take on the values 0 and 1. The random variable L_i takes on the value 0 only when $z_{2i-1} = 0$ and $z_{2i} = 0$. Since z_{2i-1} and z_{2i} are independent, we may write

$$P(L_i = 0| H_0) = P(z_{2i-1} = 0| H_0) P(z_{2i} = 0| H_0) \tag{C.2}$$

$$P(L_i = 0| H_1) = P(z_{2i-1} = 0| H_1) P(z_{2i} = 0| H_1) \tag{C.3}$$

By making use of (5.2) and (5.3), the above two probabilities become

$$P(L_i = 0| H_0) = (1 - PFL_{2i-1}) (1 - PFL_{2i}) \tag{C.4}$$

$$P(L_i = 0| H_1) = (1 - PDL_{2i-1}) (1 - PDL_{2i}) \tag{C.5}$$

In a similar manner, we can obtain the rest of the values that L_i assumes alongwith their probabilities under H_0 and H_1. These are shown in Table C1 below.

L_i	$P(L_i|H_0)$	$P(L_i|H_1)$
0	$(1-PFL_{2i-1})(1-PFL_{2i})$	$(1-PDL_{2i-1})(1-PDL_{2i})$
A_{2i-1}	$PFL_{2i-1}(1-PFL_{2i})$	$PDL_{2i-1}(1-PDL_{2i})$
A_{2i}	$(1-PFL_{2i-1})PFL_{2i}$	$(1-PDL_{2i-1})PDL_{2i}$
$A_{2i-1}+A_{2i}$	$PFL_{2i-1}PFL_{2i}$	$PDL_{2i-1}PDL_{2i}$

Table C1 : Probability distribution of L_i

The values that L_i assumes arranged in an ascending order are shown in the figure below.

By examining (5.31), we see that both A_{2i-1} and A_{2i} are independent of the regional threshold η_i while the equivalent threshold C_i is a function of η_i. Since at the moment there are no restrictions on the values assumed by η_i, then this means that C_i also may assume any value. Therefore, the possible values of L_i partition the space of C_i into five regions. If the threshold C_i lies in Z_{i0}, then according to (5.33), $PFR_i = PDR_i = 1$ (here all the values of L_i are greater than C_i and, therefore, $P\{L_i \geq C_i\} = 1$). This also means that $1-PFR_i = 1-PDR_i = 0$. Let us assume that regional detector RDi is the only detector with $PFR_i = PDR_i = 1$. All other regional detectors may have other values of the false alarm and detection probabilities. Substituting these probabilities into (5.4) results in the value of the product term to be equal to 0. This means that the value of the second term on the right hand side of equation (5.5) is 0. Therefore, under these conditions the value of the MAC

,R, is equal to R_0 which is the maximum possible value of R. This undesirable situation occurs when $C_i \le 0$. Using the definition of C_i in (5.31), we obtain

$$\log \eta_i \frac{(1 - PFL_{2i-1})\ (1 - PFL_{2i})}{(1 - PDL_{2i-1})\ (1 - PDL_{2i})} \le 0 \tag{C.6}$$

Expanding the logarithm of the product of two terms and rearranging the terms in (C.6), we can express the condition $C_i \le 0$ in terms of η_i as

$$\eta_i \le \frac{(1 - PDL_{2i-1})\ (1 - PDL_{2i})}{(1 - PFL_{2i-1})\ (1 - PFL_{2i})} = D_{i1} \tag{C.7}$$

Similarly, if C_i is such that $C_i \in Z_{i4}$, then $PFR_i = PDR_i = 0$ (in this case all the values of L_i are smaller than C_i and, therefore, $P\{L_i \le C_i\} = 0$. Using these probabilities in evaluating (5.4), we see that the product term is 0 and that the second term on the right hand side of (5.5) is also 0. This situation will lead to the result that the global cost $R = R_0$ which, of course, is not a desirable result. The values of C_i that lead to this undesirable situation are given by $C_i > A_{2i-1} + A_{2i}$. This condition can be expressed as

$$\log \eta_i \frac{(1 - PFL_{2i-1})\ (1 - PFL_{2i})}{(1 - PDL_{2i-1})\ (1 - PDL_{2i})} > A_{2i-1} + A_{2i} \tag{C.8}$$

From the definitions of A_{2i-1} and A_{2i} in (5.31), we can simplify (C.8) so that we get the following undesired values of η_i

$$\eta_i > \frac{PDL_{2i-1} PDL_{2i}}{PFL_{2i-1} PFL_{2i}} = D_{i4} \tag{C.9}$$

Since $\eta_i \le D_{i1}$ and $\eta_i > D_{i4}$ result in the maximum possible value of the MAC, η_i should not assume values in these regions. If η_i corresponds to the values of C_i such that

$$0 < C_i \leq \min (A_{2i-1}, A_{2i}) \tag{C.10}$$

then the system MAC remains the same for all possible values of η_i in this region. This is due to the fact that PFR_i and PDR_i remain the same in this region. Substituting into (C.10) the value of C_i given in (5.31) we get

$$0 < \log \eta_i + \log \frac{(1 - PFL_{2i-1})\,(1 - PFL_{2i})}{(1 - PDL_{2i-1})\,(1 - PDL_{2i})} \leq \min (A_{2i-1}, A_{2i}) \tag{C.11}$$

By using the definition of D_{i1} in (C.7), we can express (C.11) as

$$\log D_{i1} < \log \eta_i \leq \log D_{i1} + \min (A_{2i-1}, A_{2i}) \tag{C.12}$$

Therefore, In terms of η_i, the region defined by (C.10) can be expressed as

$$D_{i1} < \eta_i \leq D_{i2} \tag{C.13}$$

where

$$D_{i2} = D_{i1} exp \,(\min (A_{2i-1}, A_{2i}))$$

The two remaining regions of η_i are obtained from the following regions of C_i

$$\min (A_{2i-1}, A_{2i}) < C_i \leq \max (A_{2i-1}, A_{2i}) \tag{C.14}$$

$$\max (A_{2i-1}, A_{2i}) < C_i \leq A_{2i-1} + A_{2i} \tag{C.15}$$

Similar to what we did above, these two regions can be expressed in terms of η_i as

$$D_{i2} < \eta_i \leq D_{i1} exp \,[\max (A_{2i-1}, A_{2i})] \ = D_{i3} \tag{C.16}$$

$$D_{i3} < \eta_i \leq D_{i4} \tag{C.17}$$

The above partitions for η_i can be used in the optimization process.

Appendix D

In this appendix we prove the inequalities given by (6.32) and (6.50). We begin with the inequality given (6.32). This inequality can be verified as follows. Assume for the moment that (6.32) is true. This inequality can be written as

$$\frac{\Delta P_D}{(1 - P_D)} > \frac{\Delta P_F}{(1 - P_F)} \qquad \text{(D.1)}$$

Integrating each side of (D.1) with respect to the variable involved we get

$$\int_0^{P_D} \frac{dP_D}{(1 - P_D)} > \int_0^{P_F} \frac{dP_F}{(1 - P_F)} \qquad \text{(D.2)}$$

Performing the integration in (D.2) we get

$$-ln\,(1 - P_D) > -ln\,(1 - P_F) \qquad \text{(D.3)}$$

This inequality can be written as

$$ln\,(1 - P_F) > ln\,(1 - P_D) \qquad \text{(D.4)}$$

Since the logarithmic function ln (t) is an increasing function of t, then from (D.4) we readily obtain the following inequality

$$P_D > P_F \qquad \text{(D.5)}$$

The inequality in (D.5) is known to be true because of the concave nature of the receiver

operating characteristic of the local detectors indicating that the inequality given by (6.32) is true indeed

Next, we prove the inequality given by (6.50). This inequality can be verified as follows. First assume that (6.50) is true and rewrite it in the equivalent form

$$\frac{P_D}{P_F} > \frac{\Delta P_D}{\Delta P_F} \tag{D.6}$$

The right hand side of (D.6) represents the slope of the receiver operating characteristic of the local detector, while the left hand side represents the ratio of the detection to the false alarm probabilities at the point where we evaluate the slope. Since the receiver operating characteristic is a concave function, it follows that the curve representing the slope lies above the curve representing the ratio of the detection to the false alarm probabilities for all points. This is clear from Fig. 6.2. It follows immediately that (D.6) is true, and therefore our initial assumption regarding (6.50) is also true.

Biographical Data

NAME OF AUTHOR : Wael Abdel-Mohsen Hashlamoun

PLACE OF BIRTH :   Hebron, Palestine

DATE OF BIRTH :   June 9,1957

GRADUATE AND UNDERGRADUATE SCHOOL ATTENDED :

University of Jordan, Amman, Jordan

Syracuse University, Syracuse, New York

DEGREES AWARDED

Master of Science in Electrical Engineering, 1983,
Syracuse University

Bachelor of Science in Electrical Engineering, 1981
University of Jordan, Amman, Jordan

PROFESSIONAL EXPERIENCE:

Teaching Assistant, Department of Electrical Engineering,
Birzeit University, 1981-1982

Lecturer in the Department of Electrical Engineering,
Birzeit University, 1983-1987

Teaching and Research Assistant, Department of Electrical
and Computer Engineering, Syracuse University 1987-1991

MISSION

OF

ROME LABORATORY

Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence (C^3I) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.